The Dissertation Committee for Jiameng Zheng
certifies that this is the approved version of the following dissertation:

Three Essays on the Economics of Water Pollution Control

Committee:

Sheila Olmstead, Supervisor

Raissa Fabregas robles gil

Lynn Katz

Andrew Waxman

Three Essays on the Economics of Water Pollution Control

by

Jiameng Zheng

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2021

Dedicated to my parents Ying and Jinfu, my husband Guang and my baby girl Luna.

Acknowledgments

I would like to express my deepest appreciation to my advisor, **Dr. Sheila Olmstead**, for her relentless support and invaluable advice during my study. It is truly my great honor and luck to have such an amazing mentor. I wouldn't be able to start my Ph.D. in 2016 without Sheila's support, to begin with. Since I moved to Austin, Sheila has given me multiple opportunities to work alongside her and learn how to become an independent researcher through many research projects that shape this dissertation. When I start working on my single-authored job market paper, Sheila helped my application for multiple restrictive data sets and external grants, offered me invaluable insights and constructive advice weekly, and even helped me with edits. Moreover, when I don't have confidence in myself and my research, Sheila always believes me and encourages me to keep working on my projects. I experienced some important life-changing events during my Ph.D., such as when I lost my mother-in-law to cancer, when I had my baby girl Luna, when I was trying to get on the job market last minute, Sheila was always there to support me and guide me through those tough times. The completion of my dissertation would not have been possible without Sheila.

I'm also extremely grateful to my other committee members, **Dr. Andrew Waxman, Dr. Lynn Katz, and Dr. Raissa Fabregas**, for their valuable suggestions and guidance on my dissertation. Andrew and Raissa helped me think through my problem and identify potential models for my analysis when I was struggling to start my job market paper, offered me practical suggestions and invaluable insights on how to write the paper and survive the job market, and guided me through the whole process. When I lose confidence from the mentally draining job market, they encouraged me and helped me write a better paper and be a better version of myself. I also want to thank Lynn for her insights on the dissertation. As someone without any background in chemistry, I learned tremendously from her class and her comments. I won't be able to finish my job market paper and feel confident with the science without her insights.

I would also like to extend my deepest gratitude **Dr. Yusuke Kuwayama**, who though not on my committee, provided me with helpful advice and unwavering guidance throughout my Ph.D. Yusuke guided me and shaped the hedonic paper when Sheila was in the White House. He also provided me many constructive advice, practical suggestions, and supports when I was on the job market.

I am also grateful to many great professors I met during my study at the University of Texas at Austin, such as **Prof. Victoria Rodrigez, Peter Ward, Chandler Stolp, Bill Spelman, Cynthia Osborn, Paul von Hippel, Sandra Black** and many other professors provide me with the tools that I needed to choose the right direction and complete my dissertation.

Many thanks to my friends, **Xue Gao, Eun Young Kim, Fran-**

cisca Bogolasky, Selena Caldera, Cale Reeves, Ashlyn and Mark Hand, Alfonso Rojas-Alvarez, Ana Canedo, Noorah Alhasan, Diana Bolsinger, Regina Buono, Maria-Elena Giner, Julian Plough, Sophie Morse, Vivek S. Shastry, Santiago Tellez, Noah Durst, Ilse Oehler and many others. I feel lucky to have a supportive group of friends that provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

Last, my deep and sincere appreciation to my family, Dianfa, Chunjie, Ying, Jinfu, Guang, and Luna. It was their unconditional trust, timely encouragement, and endless patience that enable me to finish this journal. I am forever in the debt of my parents for giving me the opportunities and experiences that make me who I am and also the supports for pursuing my dream. Guang and Luna have been extremely supportive and patient throughout the process, especially as we finished a cross-country move as I'm finishing up the dissertation. This dissertation won't be possible without them and I dedicate this milestone to my family.

Three Essays on the Economics of Water Pollution Control

Publication No. ____________________

Jiameng Zheng, Ph.D.
The University of Texas at Austin, 2021

Supervisor: Sheila Olmstead

Water pollution poses important challenges worldwide. In developed countries, most of the challenges from water pollution have to do with recreational and amenity use of water, as well as the negative impact on ecosystems. For instance, in the United States, dead zones caused by nutrient pollution occur annually in many major coastal waters, including Tampa Bay, the Gulf of Mexico, Chesapeake Bay, and coastal North Carolina, causing large welfare effects in these regions. In developed countries like the United States, the aging drinking water infrastructure, such as the presence of lead pipes, is also a threat to human health. In developing countries, water pollution has a pronounced impact on human health given that safe drinking water is limited in many areas.

Economic analysis plays a critical role in the making of environmental policy. In designing and assessing a water pollution control policy, it is

important to understand the costs and benefits of such policies and be able to empirically evaluate their effectiveness. However, there are still important challenges in understanding the costs and benefits of water pollution control policies. Water quality improvement is a non-market good, so no direct price signal is available for valuing it. To overcome this problem, economists have developed several non-market valuation techniques, such as hedonic property models and recreation demand models. Each valuation method only captures a piece of the price consumers are willing to pay to improve water quality.

This dissertation comprises three papers that answer some critical questions on the economic analysis of water pollution policies. In the first paper, I estimate the marginal willingness-to-pay of homeowners for water quality improvement in Florida,using a two-stage model that combines the recreational value and amenity value of both local and regional water quality improvement. This paper, which focuses on nutrient pollution problems related to the dead zones discussed earlier, generates a more comprehensive estimate of the benefits of water pollution reduction than that used in prior work. In the second paper, I estimate an important cost of water pollution by investigating the short-run and long-run educational impacts of lead pollution in drinking water. Using data from Texas, I find that drinking water lead exposure at birth has a significant negative impact on both 3rd-grade standardized test scores and the high school graduation rate. While many prior papers in environmental economics quantify short-run and long-run human capital costs of air pollution, this paper is one of only a few to do so for an important water pollution problem. Switching to the third paper, I examine the existing literature on the policy instruments that can be used to reduce water pollution. With a focus on developing countries, I describe the empirical evidence on the effectiveness of various water pollution control policies, identify the challenges for implementing and assessing such policies, and provide recommendations for future research.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Water pollution poses important challenges worldwide. In developed countries, most of the challenges from water pollution have to do with recreational and amenity use of water, as well as the negative impact on ecosystems. For instance, dead zones are a phenomenon in which the amount of oxygen dissolved in water become too low to support many aquatic organisms, due to excess loading of nutrients like Nitrogen and Phosphorous from farms, and other sources. The occurrence of coastal dead zones has increased dramatically in recent decades, doubling to over 400 zones from 1995 to 2008 and increased to 515 sites in 2011(Rabotyagov et al. 2014). In the United States, dead zones occur annually in many major coastal waters, including Tampa Bay, the Gulf of Mexico, Chesapeake Bay, and coastal North Carolina, causing large welfare effects in these regions (Massey et al. 2006). In developing countries, water pollution has a pronounced impact on human health given that safe drinking water is limited in many areas. For example, Ebenstein (2012) finds that one-grade degradation of water quality on a six-grade scale is associated with a 9.6% increase in the digestive cancer death rate in China.

The rapid development of environmental economics over the past 20 years provides economists theoretical bases and methodological tools to estimate the benefits and costs of alternative courses of action (Morgenstern 2014). Economic analysis plays a critical role in the making of environmental policy. Water quality is a non-market good, so no direct price signal is available for valuing pollution control. To overcome this problem, economists have developed several non-market valuation techniques, such as hedonic price models and recreational demand models. Each valuation method only captures a piece of the price consumers are willing to pay to improve water quality.

Governments enact water pollution control policies to reduce water pollution, which may improve water quality. Economists use water quality parameters in economic analyses to evaluate the effectiveness of water pollution control policies and to estimate their benefits and costs,to inform the future policy-making process. Figure 1.1 describes this process and also serves as a road map of this dissertation.

This dissertation includes three papers that answer some critical questions on the economic analysis of water pollution control. In the first paper, I estimate the benefits of water quality improvement. I calculate marginal willingness-to-pay of homeowners for water quality improvement in Florida by utilizing a two-stage model, combining the recreational and amenity value of both local and regional water quality improvement. This paper, which focuses on pollution problems related to the dead zones discussed earlier, generates a more comprehensive estimate of the benefits of water pollution reduction than that found in prior work. For my second paper, I focus on the cost of a different

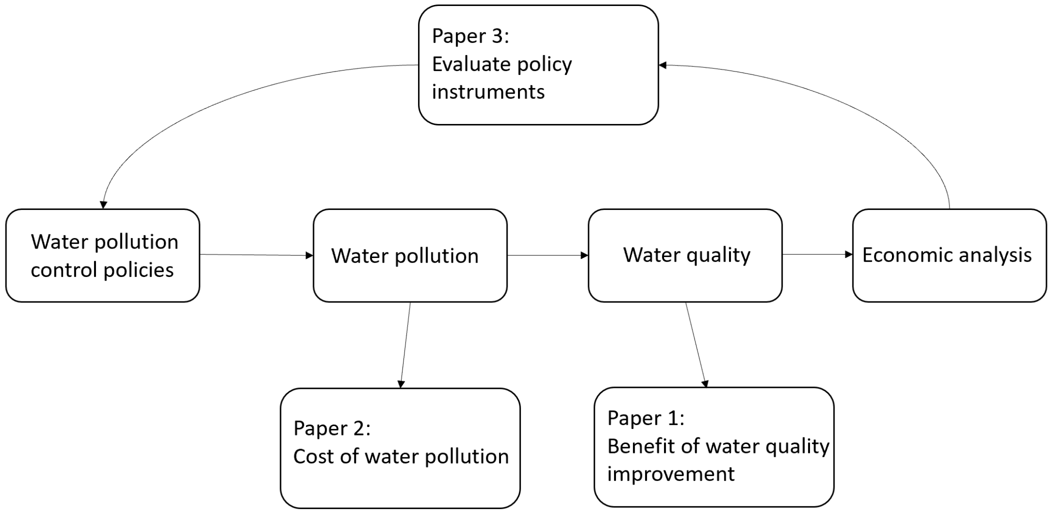


Figure 1.1: Road map of dissertation

water pollution problem, investigating the short-run and long-run educational impacts of lead pollution from drinking water. This is one of the first studies on both the short-run and long-term costs of water pollution. Switching to the third paper, I examine the existing literature on the policy instruments that can be used to reduce water pollution. With a focus on developing countries, I describe the empirical evidence on the effectiveness of various water pollution control policies, identify the challenges for implementing and assessing such policies, and provide recommendations for future research.

Paper 1: The Value of Water Quality: Separating Amenity and Recreational Benefits

Hedonic property studies that value water quality improvements generally focus on waterfront homes, or those very close to affected water bodies. The estimated marginal willingness to pay (MWTP) for pollution reduction in these studies is typically small and drops sharply with distance from the water body. One challenge with the hedonic approach is that it is unclear what these MWTP estimates capture. Unlike in the case of air pollution, health benefits from ambient water quality improvements are unlikely to be a significant share of estimated MWTP. Existing estimates likely combine primarily amenity benefits of water pollution reductions and recreational benefits. While amenity benefits may be highly localized, as prior studies have shown, recreational benefits may not be, and prior hedonic work may have failed to capture the potentially significant influence of recreation on MWTP for water quality improvements. Using the case of nutrient pollution reductions in Tampa Bay, Florida, we estimate a two-stage model that combines a random-utility recreational demand model with a hedonic housing model, allowing households to optimize over regional aquatic recreation opportunities (influenced by pollution in recreational waters), as well as local ambient water quality very close to homes. Preliminary results indicate that Tampa homeowners exhibit significant MWTP for both improvements in local ambient water quality, and improvements in regional recreational waters.

This paper makes several contributions to the literature. First, using repeat-sales from 1998-2014, this paper is the first to use a property fixed-effects model to estimate the WTP for water pollution control. This is an important contribution, given that we are able to control flexibly and comprehensively for non-time-varying property characteristics. A second contribution is that our approach allows us to expand the spatial extent of water quality impacts on property prices. Our approach, by integrating a random utility model with a hedonic model, allows us to estimate the value of recreational water quality improvements across the whole metro area housing market, rather than valuing water quality benefits to only homeowners near water bodies or visitors only at recreation sites.

This paper is coauthored with Dr. Sheila Olmstead and Dr. Yusuke Kuwayama. Kuwayama, Olmstead, and I designed the study, I cleaned the data and performed the analysis, and Kuwayama, Olmstead, and I wrote the manuscript.

Paper 2: The impacts of drinking water lead exposure on short-run and long-run educational outcomes

In this paper, I estimate the short-run and long-run impacts of early childhood lead exposure from drinking water on educational outcomes, the spatial and demographic distribution of these impacts, and the welfare effects of lead abatement policies. Using data from the U.S. Environmental Protection Agency on lead violations under the Safe Drinking Water Act and data on individual standardized test scores and educational attainment from restrictive–use data from Texas, I use two instrumental variables (IV) strategies and a difference-in-difference model to plausibly causally estimate the impact of lead. I find that lead exposure at birth has a significant negative impact on students' 3rd-grade standardized test scores and on students' ability to pass standardized tests. In the long run, experiencing a drinking water lead treatment violation at birth also significantly reduces high school graduation rates. Moreover, I find that drinking water lead disproportionately affects female children, children from African American families, and families with economic disadvantages. These results contribute to a growing literature documenting substantial long-term consequences of lead exposure and its contribution to inequality.

This paper makes the following contributions to the literature. First, I provide the first evidence of the impacts of contemporary U.S. drinking water lead level on elementary school test scores, showing that exposure even at the low levels typical of regulated U.S. water systems may cause damages. Most papers that estimate damages from lead exposure focus on airborne lead from gasoline, few studies examine the impact of lead exposure from drinking water, and most of those consider historical exposure at much high levels, where lead pipes were more common in the United States.

Second, this paper provides the first evidence that these effects persist through longer educational milestones, such as high school graduation. Previous studies have documented lead's impact on children's early cognitive ability, intelligence score (Ferrie et al. 2012), educational outcomes (Reyes 2015*b*, Aizer & Currie 2019), and later crime rate, risky behavior (Reyes 2015*b*) and juvenile delinquency (Aizer & Currie 2019). However, few studies estimate the long-run impacts of lead exposure, such as the impact on higher educational attainment or future earnings.

Third, this paper also contributes to the literature on lead exposure's implications for inequality. Economic and racial inequality can cause poor and minority children to have greater exposure to lead, and prior work suggests that lead may be one cause of continuing disparities in test scores (Aizer & Currie 2019). Following the resurgence of the Environmental Justice (EJ) literature, understanding the causes of inequality in lead exposure and consequences of lead exposure for social inequality could contribute to the discussion on new approaches and policies for reducing inequality (Banzhaf et al. 2019).

Fourth, this paper also proposes a new instrumental variable to characterize potentially endogenous drinking water lead exposure. For this purpose, I develop a novel instrument based on water chemistry– the surface source water chloride concentration – that has not been applied previously in the economics literature.

Paper 3: Water Pollution Control in Developing Countries: Policy Instruments and Empirical Evidence

Severe ambient water pollution is common in many developing countries. A broad array of regulatory and other policy instruments can be used to improve water quality. However, some approaches have been studied more than others, and there are many additional challenges that are specific to the developing country setting. This article describes a range of prescriptive and market-based regulations, voluntary programs, and other policy instruments to control water pollution and reviews the empirical evidence on the effectiveness of these approaches in practice, with a focus on developing countries. We also examine additional challenges for implementing and assessing such policies in developing countries, including data availability and quality issues, in-sufficient monitoring and enforcement, rent-seeking in regulatory systems, and jurisdictional spillovers where regulation is decentralized. Finally, we highlight important gaps in the published empirical research in this area.

This paper is co-authored with Dr. Sheila Olmstead. Olmstead and I designed the study, performed the literature review, and wrote the manuscript together.

Chapter 2

A More Comprehensive Estimate of the Value of Water Quality

2.1 Introduction

Valuation of non-market environmental amenities such as clean air and water is a long-standing challenge in economics. Revealed preference approaches tend to be preferred over stated preference approaches, and the literature (especially for air pollution) has developed significantly over the past few decades. The hedonic property model, a prominent valuation tool attributed to Rosen (1974), monetizes pollution and pollution control impacts via their influence on property prices.

Plausibly causal estimates of the value of environmental amenities and disamenities using hedonics have valued proximity to hazardous waste sites (Greenstone & Gallagher 2008), shale gas wells (Muehlenbachs et al. 2015), and improvements in air quality under the Clean Air Act (Bento et al. 2015, Bajari et al. 2012). The general approach in contemporary hedonics defines a circle of influence around properties in the sample – for example, assuming that air quality affects property values at some standard radius – usually performing sensitivity analysis around the baseline radius and reporting a range of results.

This seems, intuitively, to be good practice when household members tend to be exposed to the environmental condition primarily at or near their home.

Water pollution has also been valued using hedonic property methods using this approach (e.g. Keiser & Shapiro (2019*b*)). The assumption that exposure to water pollution occurs primarily at or near one's property may not be tenable, however. In this paper, we argue for a departure from the long prior literature using hedonics to value water quality changes, based on this premise. Our basic intuition is that, while property owners likely have some marginal willingness to pay (MWTP) for pollution reductions in small creeks, canals, streams, ponds, lakes and other waterbodies near their homes, their MWTP for water quality is likely also influenced by the degree to which water quality affects regional recreational opportunities. For example, a resident of Brooklyn, New York, may value improvements in water quality in the Gowanus Canal if they live nearby; the canal may smell better and be more visually appealing, for example. But Brooklyn residents may also value improvements in water quality at Brighton or Rockaway Beaches, or the fact that they can compete in the New York City triathlon with a swim portion in the Hudson River. However, the standard hedonic approach may not capture such benefits of improvements in waterbodies that are farther away from the homes of these residents. A more comprehensive economic valuation framework is needed in order to evaluate the benefits of major water quality improvements and compare them with costs.

In this paper, we apply such a framework to the case of water pollution abatement in Tampa Bay, Florida. We show that the recreation benefits of reducing water pollution are substantial, and that excluding them results in dramatic underestimation of benefits. We demonstrate that the hedonic property model may not be well-suited in its basic formulation to capture the recreational benefits of water pollution abatement. The standard hedonic property approach fails in this setting because, unlike air pollution, individuals in high-income countries like the United States are exposed to ambient water pollution via recreation at times and in places of their choice, at locations that may be some distance from where they live. Thus, an accurate estimate of water pollution abatement benefits at recreation sites requires an approach that matches property owners with the sites they frequent.

Theoretically, we are motivated by an integrated, two-part model of recreation and housing demand developed by Phaneuf et al. (2008). The first stage consists of a random-utility model of recreation demand, with which we estimate Tampa Bay households' indirect utility from recreational fishing trips. The second stage involves a hedonic property model. The time-varying independent variables in the hedonic model include both local ambient water quality very close to each home and estimates of indirect utility from the first-stage recreation demand model, such that our hedonic estimates of MWTP reflect the value of both amenity and recreational improvements due to water pollution abatement. As a result, we are able to estimate separately the portions of property value increases due to water quality improvements that can be attributed to amenity and recreational benefits. In their original application, Phaneuf et al. (2008) used cross-sectional property price and water quality data, and assessed recreation behavior with a household survey. We adapt this model in two ways. First, we adopt a panel data approach and exploit variation from repeat home sales using both property FE models and an innovative long-difference hedonic model. Second, we match households with recreation behavior using a national recreational fishing survey in which visitation is captured at the zip code level, rather than at the household level. The first adaptation is an improvement over prior work with respect to identification. The second is done out of necessity; because household surveys are costly, the approach we take to dealing with measurement error in attributing zip-code-level average recreation behavior to individual households may be useful in other applications.

The water pollution problem we examine in Tampa Bay is nutrient over-enrichment and eutrophication, a common water quality problem, especially in coastal areas. During our study period, 1998–2014, the region successfully reduced nutrient pollution in the watershed and experienced notable improvements in water quality. We find significant household MWTP for these nutrient pollution reductions driven by both local amenity benefits and improved recreation opportunities; both factors are capitalized into housing prices and both are statistically and economically significant. Using our more conservative long-difference approach, for the observed average 10 percent increase in dissolved oxygen (our main indicator of good water quality) in the watershed from 1998 to 2014, our baseline estimates suggest that homeowners' valuation of the marginal improvement in very local water quality—an indicator of amenity values—is about \$440 per home. Applied to all the repeat-sales homes in our sample, the total value of the observed improvement in amenity values is about \$75 million; applied to all owner-occupied homes in the Tampa metro area, the aggregate value is about \$356 million. The Tampa housing market capitalized much larger values for the impact of water quality improvements on regional recreation opportunities over the same time period: about \$980 per household, which aggregates to \$167 million for our entire repeat-sales sample and \$789 million for all owner-occupied homes in the metro area. A comparison of these benefit estimates to a very rough estimate of the costs of obtaining the observed water quality improvements suggests a favorable benefit-cost ratio. Though we focus only on a single coastal city, our results suggest that omission of recreational benefits within a hedonic framework may result in dramatic underestimation of the value of water quality improvements to homeowners. More comprehensive estimates that that capture homeowner benefits from both local and regional water quality improvements—like the ones we present in this paper—may serve as counterpoints to the existing lack of evidence that the benefits of water quality exceed the billions of dollars that are spent controlling water pollution in the United States every year.

The rest of the paper proceeds as follows. In Section 2.2, we review the prior literature on the economic benefits of water quality improvements. Section 2.3 presents our theoretical model. Our data and study area are described in Section 3.5, and econometric models are presented in Section 2.5.

Section 3.6 summarizes the main results and robustness checks. In Section 3.7, we implement our rough benefit-cost analysis and we conclude.

2.2 Literature Review

2.2.1 Hedonic Analysis and Water Quality

Many hedonic analyses in the prior literature estimate the impacts of one or more water quality parameters on property prices, starting with Epp & Al-Ani (1979) and continuing through Keiser & Shapiro (2019*b*).[1] While all of these published studies find significant positive effects of water quality improvements or, conversely, negative effects of water pollution, on property prices, only one (Mendelsohn et al. 1992) uses property fixed effects to control flexibly and comprehensively for non-time-varying property characteristics.[2] Omitted variables are a significant concern in these analyses, given the likely correlation between unobserved drivers of property prices (such as proximity to areas with high runoff or point source emissions) and water pollution. Some of the later papers in this literature use neighborhood fixed effects and difference-in-differences approaches (Horsch & Lewis 2009, Keiser & Shapiro

[1]The full list of hedonic analyses includes: Epp & Al-Ani (1979), Young (1984), d'Arge & Shogren (1989), Mendelsohn et al. (1992), Steinnes (1992), Boyle et al. (1999), Leggett & Bockstael (2000), Poor et al. (2001), Gibbs et al. (2002), Boyle & Bouchard (2003), Poor et al. (2007), Phaneuf et al. (2008), Horsch & Lewis (2009), Zhang et al. (2010), Walsh et al. (2011), Netusil et al. (2014), Wolf & Klaiber (2017), Walsh et al. (2017), Keiser & Shapiro (2019*b*).

[2]Theoretically, hedonic analysis involves two stages. The first stage is the use of property prices and characteristics to obtain a marginal implicit price. The second stage estimates a demand curve for the environmental good or service to use for welfare analysis. Given limitations in data availability, most empirical analyses focus on the first stage.

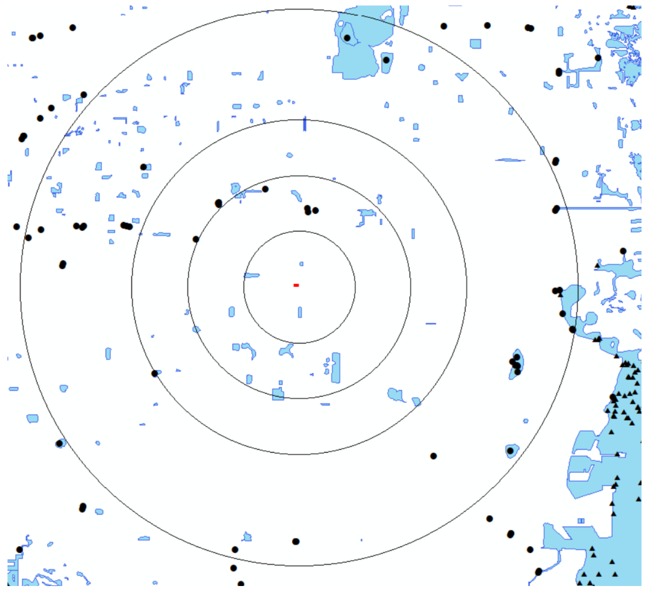
2019*b*) that likely provide a good approximation to models in which the identifying variation comes from repeat sales of the same property over time. However, given that the repeat-sales approach has become standard in the hedonics literature for non-water-quality applications (Bajari et al. 2012, Muehlenbachs et al. 2015, Walls et al. 2015), applying this approach in an analysis of water pollution is one of our significant contributions.

A second critical difference between our paper and prior work is our approach to defining the spatial extent of water quality impacts on property prices. Because the benefits of water quality improvements may vary with distance to the waterbody, some papers only attempt to quantify impacts on waterfront homes (Leggett & Bockstael 2000, Poor et al. 2001, Gibbs et al. 2002, Zhang et al. 2010). Other papers estimate benefits at different distances from the waterbody (Poor et al. 2007, Walsh et al. 2011, 2017, Guignet et al. 2017, Keiser & Shapiro 2019*b*). In this literature, MWTP for water quality diminishes quickly with distance, generally between 2 and 3 kilometers (km) from the water.

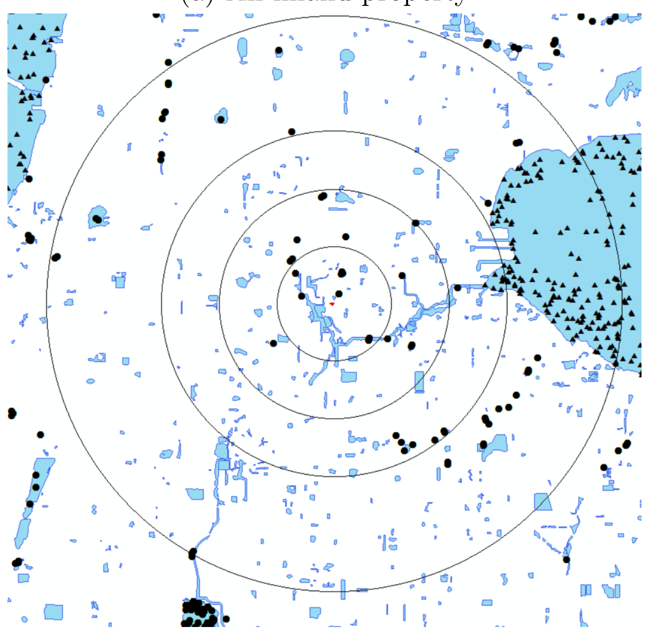
Our concern with these findings is that they may not fully capture the recreational benefits of water quality. Most individuals do not recreate in waters within 3 km of their property. For example, in our sample, recreational anglers' average roundtrip travel time to fishing sites is almost 90 minutes. Unlike in the case of air quality, where health impacts occur everywhere individuals spend time (e.g., at home or on a commute), individuals generally choose when and where they recreate in or near water (and thus experience water pollution), and examining impacts based on simple spatial criteria of proximity of houses to water may be misleading.

In theory, hedonic property studies like those cited above could pick up both amenity and recreational benefits of water quality improvements. When economists have looked for effects outside of a very tight radius around properties, however, many have not found such effects (Walsh et al. 2011, Keiser & Shapiro 2019*b*). This has been interpreted as evidence that homeowners only value water quality close to their homes. However, the maps in Figure 2.1 demonstrate our concern about this interpretation. The red dots at the center of each panel in the figure represent two households in our study area. The house at the center of Figure 2.1a is an inland property, and the house in Figure 2.1b is located near Tampa Bay. The black dots and triangles indicate the location of water quality monitors. Concentric circles are drawn with radii of 1, 2, 3, and 5 km from the property at the center of each panel. First, consider the property in Figure 2.1a. A 2-km circle captures a handful of water quality monitors that, when averaged, may yield a reasonably good representation of water quality very close to the home. If one wanted to capture water quality in waterbodies that this household could use for recreational purposes, drawing increasingly larger circles nets many additional water quality monitors but few, if any, are in locations with which the household has any regular contact. Thus, increasing the assumed "zone of influence" for this household by drawing larger circles will attenuate any impact of willingness to pay for nearby water quality, and it will not capture recreational values.

In a coastal metropolitan area like Tampa Bay, averaging observations from water quality monitors in larger circles for properties closer to the coast like that in Figure 2.1b will capture some recreation sites, as the 5-km circle does in Figure 2.1b. However, the ability of an econometric model to effectively detect the signal of recreational water quality values from the monitors in the Bay (to the east of the property) will depend on how many irrelevant monitors (i.e., those inland and quite far from the home) are also captured. In addition, the standard hedonic model does not capture households' actual recreation sites—it only links homes to sites by proximity. Given these challenges, it is not surprising that regressing housing prices on average measures of water quality within circles around properties frequently generates null results beyond 2 km.



(a) An inland property



(b) A property near Tampa Bay

Figure 2.1: Two sample properties in Pinellas County

Notes: The red polygons in Panel A and Panel B indicate two properties in Pinellas County. The black dots are local water quality monitors, and the black triangles are recreational water quality monitors in Tampa Bay. The radii of the four circles are 1km, 2km, 3km and 5 km.

2.2.2 Recreation Demand Analysis and Water Quality

Another commonly used approach to estimate water quality benefits is recreation demand estimation using random utility models (RUMs). We identified 11 papers in the literature that use these models to value water quality changes (Mullen & Menz 1985, Smith et al. 1986, Bockstael et al. 1987, 1989, Phaneuf et al. 2000, Phaneuf 2002, von Haefen 2003, Phaneuf et al. 2008, Egan et al. 2009, Abidoye et al. 2012, Abidoye & Herriges 2012). Similar to the hedonic literature, all but one of these studies finds that water quality improvements increase recreational visitation and willingness to pay, but omitted variables bias is a concern for interpreting these results (Moeltner & von Haefen 2011, Phaneuf 2013). Only two of the studies (Abidoye et al. 2012, Abidoye & Herriges 2012) control comprehensively for both site and visitor characteristics. Three additional papers control for either unobserved site characteristics (Phaneuf et al. 2008) or visitor characteristics (von Haefen 2003, Egan et al. 2009), but not both.

The recreation demand component of our two-stage approach breaks no new ground. Rather, our contributions lie in: (1) a focus on a large, charismatic water body (Tampa Bay) that is the locus of recreational activity in a major coastal metro area and has experienced noticeable water quality improvements over the study period; and (2) integration of a RUM with a hedonic model, which allows us to estimate the value of recreational water quality improvements across the whole metro area housing market, rather than valuing water quality benefits to visitors only at recreation sites.

2.2.3 Economic Impacts of Nutrient Pollution

Nutrient over-enrichment is caused by the addition of excess nutrients, primarily nitrogen and phosphorous, to waterbodies via agricultural and urban nonpoint source pollution, which stimulates excessive algae growth. When the algae die, they decay and deplete dissolved oxygen (Morrison & Greening 2006). Because nitrogen and phosphorous enter waterbodies over a broad catchment area, there are negative impacts not only locally in small streams but also regionally in large streams, rivers, bays and estuaries. Among the serious consequences of eutrophication are hypoxic or dead zones, in which many kinds of marine life cannot be supported. Reported dead zones worldwide doubled between 1995 and 2008 to more than 400 zones, and increased to 515 sites in 2011 (Rabotyagov et al. 2014). Other than Tampa Bay, U.S. waters that experience this phenomenon include other estuaries connected to the Gulf of Mexico, Chesapeake Bay, the Great Lakes (especially Lake Erie), Puget Sound, Long Island Sound, and the North Carolina coast. Economists have estimated significant impacts of eutrophication on commercial and recreational fisheries (Massey et al. 2006, Smith et al. 2017), though other economic damages are largely unknown (Barbier 2012).[3]

[3]Anecdotal evidence suggests that recreational impacts could be significant. In 2005, one of the years included in our study period, Florida's swimming beaches experienced almost 3,500 closures and health advisories due to high levels of bacteria caused by algal blooms, including toxic cyanobacteria blooms (Clean Water Network of Florida 2008).

2.3 Theoretical Model

We adapt the theoretical model in Phaneuf et al. (2008), which has long-run and short-run decision-making components. In the long run, consumers evaluate neighborhood and property amenities, including pollution, to choose a home. In the short run, once a location is chosen, a household allocates its resources (including time) to market goods and recreation, that is, households evaluate the benefits of outings to recreation sites conditional on residential location. Since short-run recreation decisions are affected by long-run residential location choices, we assume that when making property purchase decisions, consumers will consider each location's accessibility to recreation opportunities.

Let $x(Q)$ represent a household's utility from recreation trips, where Q measures water quality at recreation sites in the region. In addition, let p_x be the price of a trip, z be a numeraire good with price equal to 1, and $h(\mathbf{a}, q)$ be the value of housing services which is quasi-fixed in the short run and is a function of a vector of property attributes, $\mathbf{a}$, and water quality close to the home, q (which can differ from regional recreational water quality Q). The household maximizes its utility for recreation trips and market goods conditional on its income after housing expenditures. Thus, the household's short-run maximization problem is:

$$\max_{x,z} \quad U(x(Q), z|h(\mathbf{a}, q)) \quad \text{s.t.} \quad m = p_x x(Q) + z, \tag{2.1}$$

where m is household income net of the property price. Note that this model assumes that households can perceive a change in Q. For instance, if nutrient pollution results in excessive algae in a recreational waterbody, households notice a change in the color of the water, a decline in fish catch, or a beach closure. Solving the problem in (2.1) yields the household's conditional indirect utility function:

$$V = V(p_x, m, Q, q, \epsilon), \tag{2.2}$$

where ϵ captures unobserved property heterogeneity.

Suppose that water quality at recreation sites improves from Q_0 to Q_1. A welfare measure for this improvement is compensating surplus (CS), which can be described implicitly by the following equation:

$$V(p_x, m, Q_0, q) = V(p_x, m - CS(m), Q_1, q). \tag{2.3}$$

That is, CS measures the income that a household is willing to forgo to obtain the improved water quality (Kim et al. 2015).

We can estimate the indirect utility from recreation using a recreation demand model. Let $CS(Q, \epsilon)$ measure the gains to a household from visiting recreation sites in Tampa Bay with water quality Q. When households make recreation decisions, they consider potential benefits and costs from visiting each possible site. If water quality and recreation costs vary spatially, different neighborhoods will offer different potential net benefits from recreation to households located in those neighborhoods. Thus, we can model expected recreational net benefits as an attribute of location:

$$ECS(Q) = \mathbb{E}[CS(Q, \epsilon)] \tag{2.4}$$

In long-run equilibrium, housing prices should capitalize the expected benefits from recreation at a given location. Since recreation decisions are made conditional on residential location decisions, we replace the $x(Q)$ in equation (2.1) with $ECS(Q)$ as defined in equation (2.4). The long-run utility maximization problem is thus:

$$\max_{\mathbf{a},q} \quad U\left(ECS(Q), h(\mathbf{a}, q), z\right) \quad s.t. \quad m^* = p_h(\mathbf{a}, q) + p_x\tilde{x} + z. \tag{2.5}$$

Households choose a residential location such that the sum of their expected marginal benefit from recreation and from services directly available from the property (including environmental services) is equal to the marginal property purchase price.[4]

2.4 Study Area and Data

The Tampa Bay watershed (Figure 2.2) covers more than 400 square miles. It contains Florida's largest open-water estuary and second-largest metropolitan area, and is the second-largest city on the Gulf of Mexico. The Bay provides important social value through species habitat and other ecosystem services, recreational use such as boating and fishing, power plant heat exchange, and commercial ports. Our study area comprises three counties within this watershed—Hillsborough, Pinellas, and Manatee counties—in which more than 2.3 million people live. Almost 90 percent of the total employment within

[4]Our model does not incorporate the fact that local water quality q and recreational water quality Q are often correlated, an important potential extension of this work.

the three counties is located in the watershed (Tampa Bay Estuary Program & Tampa Bay Regional Planning Council 2014).

Nutrient loading to the Bay originates from a variety of sources including agricultural runoff, phosphate mining, fertilizer production, urban stormwater runoff, municipal sewage treatment discharges, industrial point sources, and atmospheric deposition from power plants (Greening et al. 2014, Sherwood et al. 2016). Paired with other aspects of urbanization (for example, construction of causeways that modified the Bay's hydrology), nutrient loadings generated between the 1950s and the 1980s caused a dramatic shift from a "clear-water, seagrass-based system" to a "turbid, phytoplankton-based system" in which blooms of harmful phytoplankton were common and macroalgae mats covered large portions of open water, tidal flats, and seawalls (Greening et al. 2014). One impact of this water quality shift was an estimated 50 percent decline in seagrass coverage, an important indicator of the health of aquatic ecosystems (Greening et al. 2014).

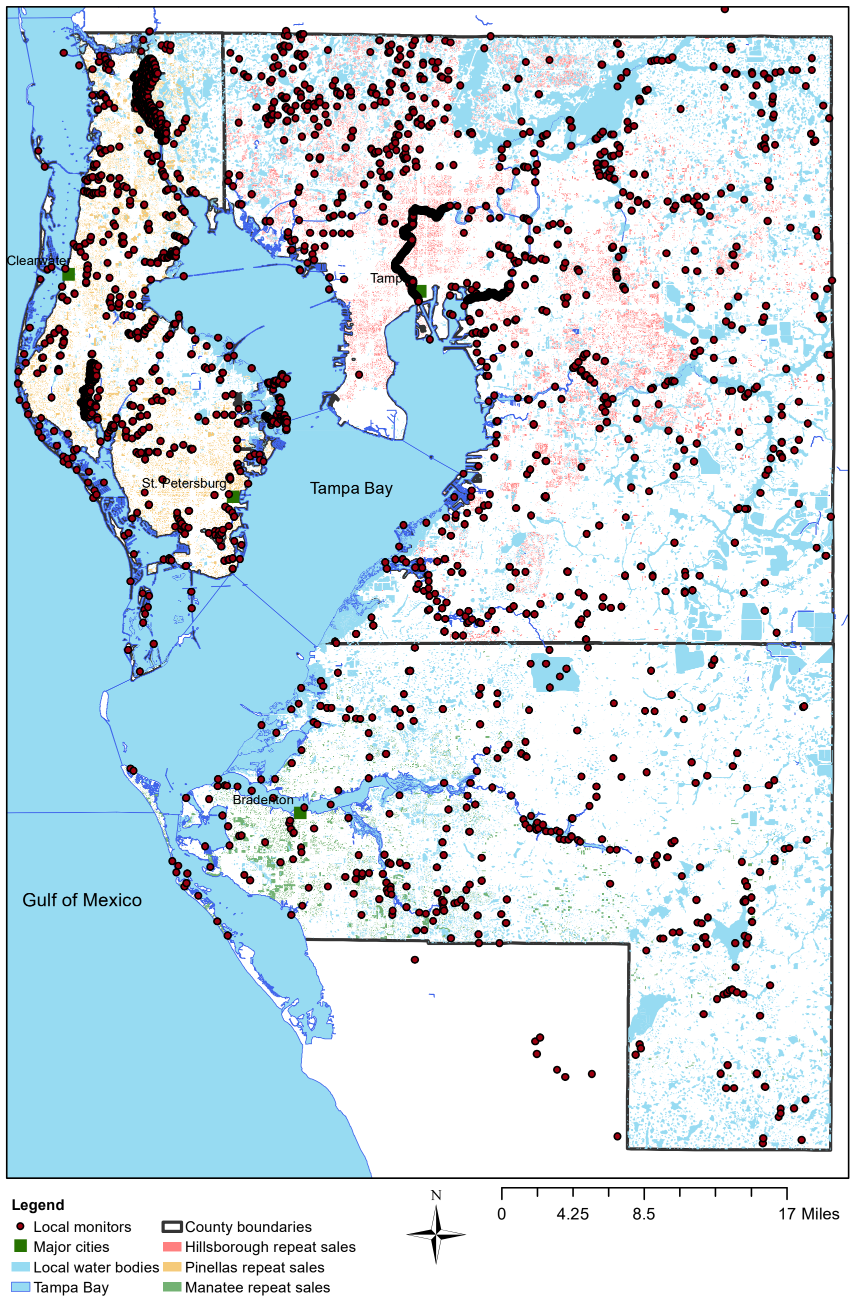


Figure 2.2: Map of study area: Tampa Bay watershed, Florida

Beginning in the late 1970s, citizens pressured the Florida legislature to impose advanced wastewater treatment standards on municipal sewage treatment plants discharging to Tampa Bay, and the resulting changes in point source emissions have been statistically associated with water quality improvements (Beck et al. 2019). This legislation was followed by new regulations and practices including statewide permitting requirements for urban stormwater systems, coastal habitat acquisition and restoration projects, fuel-switching and nitrogen oxide (NOx) abatement technology upgrades by local power plants that reduced atmospheric deposition of nitrogen, and residential fertilizer use ordinances (Beck et al. 2019). These actions led to a recovery from widespread, frequent eutrophic conditions. Tampa Bay's seagrass coverage in 2016 exceeded that observed in 1950—considered by local recovery proponents to be the "reference state" for the Bay (Greening et al. 2014). Other water quality measures are also approaching the conditions last observed in the pre-disturbance 1950s (Greening et al. 2014).

Tampa Bay's recovery and the existence of rich long-term monitoring data documenting that recovery motivate our work. While we do not observe property transactions during the entire recovery period, water quality improves noticeably over our study period, 1998–2014. In addition, the region's recreation opportunities, rapid growth and active housing market make it an ideal place for this study.

At the beginning of our study period, the U.S. Environmental Protection Agency (EPA) developed a Total Maximum Daily Load (TMDL) pollution budget for Tampa Bay covering 189 different sources, based on management targets set by the Tampa Bay Estuary Program. The Tampa Bay Nitrogen Management Council (TBNMC), a public/private partnership of local governments, agencies, and industries, developed an action plan for TMDL compliance and for supporting the Bay's continued recovery. From 1998 to 2014, the TBNMC implemented more than 600 projects to reduce nitrogen loading to the Bay (Beck et al. 2019).[5] While we cannot causally link the observed improvements in water quality to these projects, at the end of the paper, we compare our water quality benefit estimates to a rough estimate of the costs of these nutrient removal projects over our study period.

2.4.1 Recreation Demand Data

For the recreation demand model, we use angler data from the Marine Recreational Fisheries Statistics Survey (MRFSS) and the Marine Recreational Information Program (MRIP) produced by the National Ocean and Atmospheric Administration (NOAA) (NOAA Fisheries 2008). The MRIP surveys a random sample of U.S. recreational anglers (NOAA Fisheries 2013). From the MRIP, we are able to obtain the year, month and time that each interview takes place, the zip code of each angler's residential address, fishing site locations, the number of people in each fishing group, and other visit characteristics.

Since the MRIP data do not have anglers' full address or self-reported

[5]Hundreds of additional projects also preceded the TMDL.

Table 2.1: Descriptive statistics

Variable	N	Mean	Std. dev.	Min	Max
Water quality measures					
Local dissolved oxygen (DO) (mg/L)	146903	5.79	4.07	0.20	104.00
Average local DO 1998-2003 (mg/L)	14390	5.88	8.42	0.61	104.00
Average local DO 2009-2014 (mg/L)	14390	5.89	1.40	1.04	12.37
Change in average DO (mg/L)	14390	0.01	8.45	-100.22	7.99
Tampa Bay DO (mg/L)	146903	6.40	0.83	3.16	10.53
Local parks DO (mg/L)	41618	5.38	1.60	1.12	12.01
Dummy for local DO $\geq$ 5mg/L	146903	0.36	0.48	0.00	1.00
Seagrass abundance (index)	146903	2.35	0.91	0.00	4.79
Recreation demand					
Travel time (minutes)	146903	87.24	34.95	1.64	267.06
Travel cost ($)	146903	43.12	20.64	0.34	152.33
Estimated ECS ($)	146903	35.98	2.42	29.28	38.67
Average ECS 1998-2003 ($)	14390	35.30	1.56	29.28	37.40
Average ECS 2009-2014 ($)	14390	34.35	2.46	30.69	38.15
Change in average ECS ($)	14390	-0.95	2.89	-6.54	8.80
Distance to water					
Distance to Tampa Bay (m)	146903	15317.35	15212.89	0.00	120557.70
Distance to local waters (m)	146903	2796.27	1742.42	0.00	11372.92
Property characteristics					
Repeat-sales sample sale price (2014$)	146903	229306.50	154742.5	5262.23	1541511.00
Long-diff. sample sale price (2014$)	21639	166442.50	120243.80	5000.00	
Average sale price 1998-2003 (2014$)	14390	204674.30	137652.90	6873.87	1395970.00
Average sale price 2009-2014 (2014$)	14390	195782.90	140483.20	7724.32	1097957.00
Change in average price (2014$)	14390	-8891.35	63111.49	-581720.80	684031.80
Year	146903	2005.76	4.44	1998.00	2014.00
Property age (years)	146903	32.99	21.07	1.00	133.00

travel cost, we use latitude and longitude information for fishing sites and anglers' residential zip codes to estimate travel costs for each trip. We use the 2010 Census Bureau zip code tabulation area (ZCTA) maps and population data to create a population-weighted center for each zip code in the three counties using ArcGIS, and assume that all anglers live in the population-weighted center of their zip code.[6] We then use the Open Source Routing Machine API to calculate round-trip travel time from the zip code-weighted population centers to fishing sites (Luxen & Vetter 2011). Our travel cost estimate has two components: the value of this estimated travel time and the operational cost of travel. The value of travel time is estimated at 1/3 of visitors' forgone wages, using the mean hourly wage in the Tampa-St.Petersburg-Clearwater Metropolitan Statistical Area from the Occupational Employment Statistics (OES) Survey (U.S. Bureau of Labor Statistics 2018). For the operational cost of travel, we multiply the round-trip distance by the driving cost per mile reported by the American Automobile Association (AAA 2019). Both the wage and the cost-per-mile estimates vary over time. Table 2.1 reports summary statistics for our estimated travel times and costs. The mean round-trip travel time in our angler data is 87.24 minutes, or about an hour and a half, and the average trip costs $43.12.

2.4.2 Property Transaction Data

We collect property sales data from the property appraiser's offices in Hillsborough, Manatee and Pinellas Counties. In order to better identify the effect of water quality on residential property prices and maintain consistency with prior hedonic analyses, we restrict the sample to single-family

[6]Figure A.1 in the online appendix shows the locations of fishing sites, along with recreational water quality monitors and geographic boundaries in the region. The Census Bureau generates ZCTAs to represent the United States Postal Service (USPS) zip code service areas. Going forward, we refer to the US Census ZCTAs as zip codes given the fact that, in most instances, they are the same.

homes. Sales dates, dates of construction, parcel size, and transaction prices are available for all three counties. The Hillsborough County Property Appraiser's Office provided additional information, including dates of major improvements, size of living spaces, number of stories, and number of bedrooms and bathrooms. Our sample includes only homes sold at least twice between 1998 and 2014, given our desire to include property fixed effects in our hedonic specifications (Manatee County data are only available from 2005-2014). Hillsborough County has 186,289 repeat property sales that occurred during this period, Pinellas County has 107,701 repeat sales, and Manatee county has 20,699 repeat sales.[7]

We geocode the sales records and relate them in ArcGIS with shapefiles of house locations and characteristics. We then relate these property data with water quality data, also using ArcGIS. For each model we estimate, we use only properties that have at least one water quality monitor within the relevant distance and time window prior to a transaction. For example, our baseline model uses monitors within 3 km of a home to capture local water quality and uses water quality observations in the calendar year of each property transaction. Thus, for this model, we drop the 153,304 repeat sales in our data that lack water quality monitors within 3 km in the calendar year of the sale.[8]

[7]Repeat sales represent 63.2% of all sales in Hillsborough County (1998-2014), 59.7% of all sales in Pinellas County (1998-2014) and 44.9% of all sales in Manatee County (2005-2014).

[8]35.1% of repeat sales in Hillsborough (1998-2014), 62.1% of repeat sales in Pinellas

We also link properties with the zip code-year level recreational index we create, resulting in additional narrowing of the sample (in our baseline model, this requires dropping 14,482 properties). The remaining 146,903 properties comprise our full sample for the property fixed effects model—65,301 in Hillsborough County, 66,926 in Pinellas County, and 14,676 in Manatee County. The mean property price in the sample is about $230,000 (Table 2.1).[9] Properties in our sample were sold on average three times from 1998 to 2014 and were about 32 years old when a transaction occurred.

2.4.3 Water Quality Data

2.4.3.1 Local Water Quality Data

We obtained waterbody shapefiles from the Tampa Bay Water Atlas (University of South Florida Water Institute 2017), which is derived from the 1:24,000 USGS National Hydrography Dataset (NHD) and contains 749 water resources, including 12 bays, 506 lakes, 230 rivers and the Gulf of Mexico. We define ponds, lakes, wetlands, rivers, swamps, reservoirs and canals as local waterbodies and refer to water quality monitors in these waterbodies as "local water quality monitors." Water quality measures at these monitors are obtained from EPA's STOrage and RETrieval (STORET) data warehouse, which includes water quality monitoring data collected by states, tribes, water-

(1998-2014) and 70.9% of repeat sales in Manatee (2005-2014) have reporting water quality monitors within 3 km in the calendar year of the sale.

[9] All prices are in 2014 dollars. Table A.2 in the online appendix lists summary statistics for the additional property attributes available for Hillsborough County.

shed groups, federal agencies, volunteer groups, and universities. We keep all observations for which monitoring date, station latitude, and station longitude are reported. The resulting sample includes 209,336 water quality observations collected from 5,913 monitoring stations. The mean number of readings from each station per year is 53, and the monitors report on average for 8 years (see Table A.1 in the online appendix).[10]

There is no single accepted best indicator for water quality in hedonic and recreation demand analysis. Water quality measures used in past hedonic studies include dissolved oxygen (DO), fecal coliform, total suspended solids, dissolved inorganic nitrogen, pH, Secchi depth and harmful algal concentrations. We use DO, one of the most common measures of water quality in research on water pollution's economic impacts (Keiser & Shapiro 2019*b*), and a key indicator of nutrient pollution. Higher DO levels indicate better water quality. DO is critical for fish survival, and water quality that meets the criteria for fish survival also meets criteria for most other beneficial water uses and is often of good ecological status (U.S. Environmental Protection Agency 2001). DO is also a good indicator of water quality conditions that are noticed by people, and are thus likely to correlate with property prices. Noticeable impacts of low DO include reduced fish catch and the presence of algae mats.

The large red dots in Figure 2.2 depict the location of the STORET

[10]Some monitors change names slightly, and monitor identification numbers are not unique across counties. Following Keiser & Shapiro (2019*b*), we define a station as a unique latitude-longitude pair when we link properties with nearby monitors.

local water quality monitors and the smaller dots in pink, yellow, and green show the locations of repeat-sales properties in our sample. We calculate the mean DO concentration of all the monitors within varying radii in the calendar year of a property sale to generate the local water quality measure for each property. For our baseline model, we choose a 3-kilometer radius, based on existing evidence that nationwide water pollution impacts are capitalized for homes within 3 km (Keiser & Shapiro 2019*b*). We also test the robustness of our results to other radii, from 300 meters (m) to 5 km, and to varying time windows around the date of a property sale.

We also create a dummy variable indicating whether the DO level for any given observation is above 5 milligrams per liter (mg/L); a DO concentration of 5 mg/L is a critical value for fish survival and may capture a threshold for detectable impacts (U.S. Environmental Protection Agency 1994). Table 2.1 shows that the mean DO value in our sample is 5.79 mg/L, with about 36 percent of properties near waters having less than 5 mg/L DO, on average. We use the continuous DO concentration in all models in the paper, reporting the 5 mg/L threshold results in the online appendix.

Table A.2 in the online appendix lists summary statistics for properties categorized by our principal independent variable, the DO level in local water bodies. Properties near polluted water bodies are older, smaller and have fewer bedrooms, bathrooms and stories on average. They also are located further from nearby water bodies and from Tampa Bay. These differences highlight the importance of controlling comprehensively for property characteristics when estimating the impact of water pollution on property prices.

2.4.3.2 Recreational Water Quality Data

For the recreation demand model, we use DO values from STORET monitors near fishing locations in Tampa Bay, which we refer to as "recreational water quality monitors," mapped in Figure A.1 in the online appendix. Consistent with the methods we use to define water quality in local waterbodies, we spatially join all monitors within a 3 km radius of each of our 85 fishing sites and calculate the annual mean DO. The mean DO level at recreational water quality monitors is 6.4 mg/L (Table 2.1).

In addition to observations from water quality monitors, we rely on seagrass acreage measurements from the Tampa Bay Estuary Program (TBEP) (Johansson 2016). The health of Tampa Bay seagrass meadows has become an important issue in recent decades as scientists and environmental managers have worked to reverse the effects of nutrient pollution in the Bay. In 1997, the TBEP coordinated the creation of a Bay-wide seagrass monitoring program to document temporal and spatial changes in seagrass species composition, abundance, and distribution. Currently, 62 locations are monitored (Florida Fish and Wildlife Conservation Commission 2003). The TBEP's seagrass abundance data are reported as an index, with higher values representing greater abundance. We match fishing sites with their closest seagrass transects.[11]

[11] We exclude seagrass transects more than 11,000 m from each fishing site in order to avoid spatially joining fishing sites located along the west coast of Pinellas County with

Average seagrass coverage (converted from the TBEP index) is about 29,920 hectares (ha) from 1998-2014 (Table 2.1). While seagrass coverage is an important positive indicator of ecosystem health and fish abundance, it may also be a disamenity for anglers because these plants can get caught on fishing lines and boat propellers (Guignet et al. 2017), and boaters can be fined for scarring seagrass beds with their motors.

The yellow line in Figure 2.3 shows the trend in the average annual DO concentration over time, using all of the local and recreational water quality monitors in the data. While the year-to-year variation can be substantial, the trend is increasing, reflecting the regional water quality improvements described in the literature. The annual average DO concentration in 2014 is 11% higher than in 1998. Similarly, the average DO over the earliest six years (1998-2003) and latest six years (2009-2014)—the periods we will use to calculate long-run changes in some of our models—is 10%. Thus, throughout the discussion of results in Section 6, we use a 10% improvement in average DO concentrations over the 16-year study period to interpret coefficient estimates and compare them across models.

seagrass transects in Tampa Bay, which lie across the peninsula formed by Pinellas County (Figure A.1).

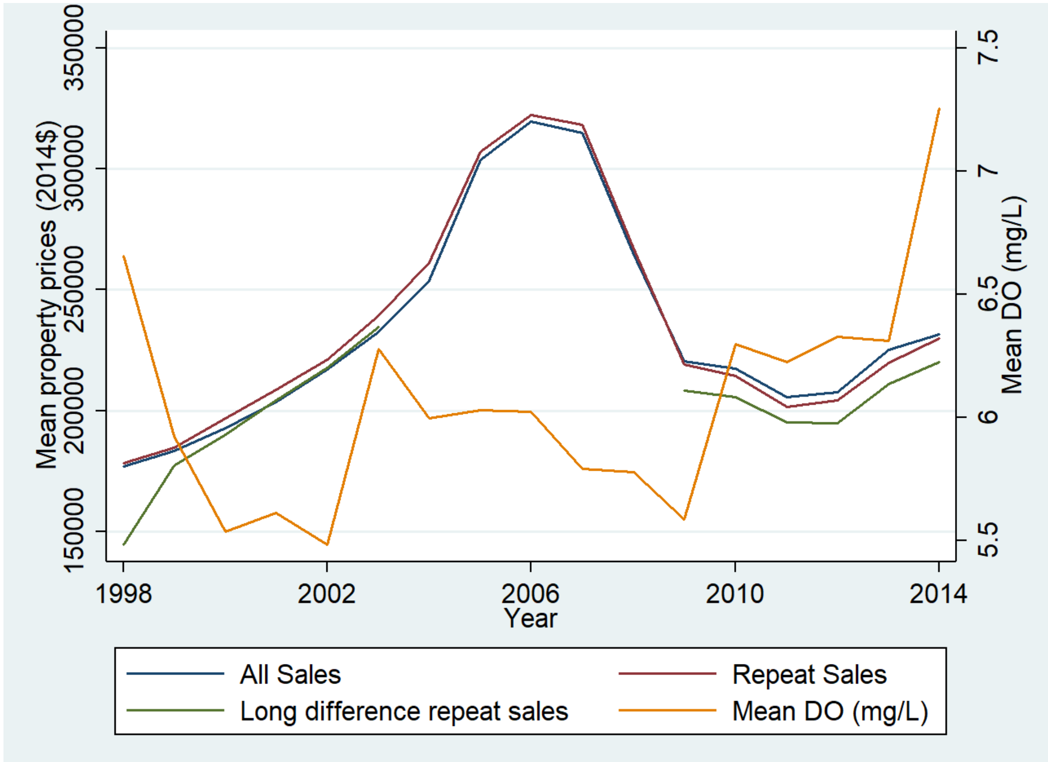


Figure 2.3: Average property prices in the three counties and average DO concentration among all water quality monitors, 1998-2014

Notes: "All sales" averages annual property transaction prices across the 241,909 properties sold in Hillsborough, Pinellas and Manatee Counties at least once, 1998-2014. "Repeat sales" averages annual prices across the 170,192 properties sold at least twice, 1998-2014. "Long difference repeat sales" averages prices across the 14,390 properties sold at least once during 1998-2003, and at least once during 2010-2014. Mean DO is the average dissolved oxygen concentration observed at all local and Tampa Bay water quality monitors each year.

2.5 Methods

2.5.1 Random Utility Specification for Recreation Demand

In the random utility model, properties in Tampa Bay are located in J zip codes, and anglers can choose to fish at K recreation sites in the region. Each recreation site $k \in K$ has an observable level of water quality, WQ_{kt}, which can vary over time. The literature recognizes the need to control for unobserved site characteristics in random utility models (Moeltner & von Haefen 2011, Phaneuf 2013). One strategy is the use of Alternative Specific Constants (ASCs)–equivalent to site fixed effects–in the basic RUM model. Following Phaneuf et al. (2008), we assume the indirect utility for a visit to site k by individual i in year t is a linear function. The RUM specification is:

$$V_{ikt} = \alpha_0 + \alpha_1 Travel_{ikt} + \alpha_2 WQ_{kt} + \eta_k + \nu_{ikt}, \tag{2.6}$$

where V_{ikt} represents indirect utility of fishing trips and $Travel_{ikt}$ denotes the round-trip travel cost. η_k is an ASC that captures time-invariant site characteristics, such as the number of boat ramps or slips, whether the fishing site has lodges, and other attributes we assume remain constant over time. We use a conditional logit model, so ν_{ikt} is an error term distributed Type-I Extreme Value.

The expected utility per trip for person i in year t is then:

$$EV_{it} = ln[\sum_{k=1}^{K} exp(\hat{V}_{ikt})] + C \tag{2.7}$$

where $\hat{V}_{ikt}$ is the observed element of utility, and C is an unknown constant indicating that the absolute level of utility cannot be measured. Because the term C in equation (2.7) does not affect utility differences, we drop it in the remaining equations (Haab & McConnell 2002). The average compensating surplus per trip is thus given by:

$$\mathbb{E}(CS)_{it} = \frac{EV_{it}}{\hat{\alpha_1}} \tag{2.8}$$

We divide EV_{it} by the coefficient on the travel cost variable, interpreted as the marginal utility of income, to obtain a monetary measure of $\mathbb{E}(CS)_{it}$. If water quality improves from WQ_0 to WQ_1, indirect utility rises from $\hat{V}^0_{ikt}$ to $\hat{V}^1_{ikt}$ (per equation (2.6)) and the average compensating surplus per trip then is given as:

$$E(CS)_{it} = \frac{1}{\hat{\alpha_1}} \left\{ ln \left[\sum_{k=1}^{K} exp(\hat{V}^1_{ikt}) \right] - ln \left[\sum_{k}^{K} exp(\hat{V}^0_{ikt}) \right] \right\} \tag{2.9}$$

Our estimate of $\mathbb{E}(CS)_{it}$ varies across zip codes and over time. The average recreational compensating surplus in zip code j in year t can be expressed as the average utility of all person-trips (N_{jt}) originating from the zip code in that year:

$$ECS_{jt} = N_{jt}^{-1} \sum_{i=1}^{N_{jt}} \mathbb{E}(CS)_{it} \tag{2.10}$$

We incorporate this estimated ECS_{jt} into our hedonic model to capture how recreational impacts of water quality improvements may be capitalized in housing prices.

2.5.2 Hedonic Specification

Our hedonic specifications control for observable and unobservable property attributes by exploiting only price changes within a property over time

(Palmquist 1982). We use two approaches: a standard property fixed-effects model and an innovative model using long differences. Further extensions to both of these basic models are discussed in Section 6.

2.5.2.1 Property Fixed Effects Model

Using a log-log specification in line with the previous literature, the basic property fixed-effects model is as follows:

$$lnP_{ijt} = \beta_0 + \beta_1 Age_{it} + \beta_2 lnWQ_{it} + \beta_3 \widehat{ECS}_{jt} + \alpha_i + \gamma_t + \epsilon_{ijt}. \tag{2.11}$$

Home age is the only time-varying property characteristic in our data. The main coefficients of interest are β_2 and β_3. Using DO as the main local water quality measure, we expect $\beta_2 > 0$ since higher DO represents better water quality. We also expect $\beta_3 > 0$ if buyers are willing to pay higher prices for properties that offer more and better recreation opportunities. We will interpret β_2 as homeowners' MWTP for an increase in DO within water bodies close to homes. We will use β_3 and equations (2.7) through (2.10) to estimate MWTP for an increase in DO in the regional recreational waters frequented by residents of homeowners' zip codes. The property fixed effect, α_i, removes the effects of time-invariant omitted variables, and we also include a year fixed effect, γ_t.

As noted in Section 4.1, the MRIP data include anglers' zip codes, but not their street addresses. Thus, our estimate of ECS_{jt} is a zip-code-level average measure of recreational utility in each year. The sample of anglers in the MRIP data in any individual household's zip code in a given year is small, and different households have different numbers of fellow anglers from the same zip code who are surveyed in the MRIP. In estimating equation (2.11), we treat the resulting potential bias from heteroskedastic measurement error as a partial missing-data problem, using multiple imputation (Blackwell et al. 2017). We first generate 250 stand-alone recreational demand datasets—representing what could have been observed if there were no measurement error—randomly drawing (with replacement) 20% of anglers (10,373 individuals) from the full MRIP sample each time. We estimate equation (2.6), and then use each set of coefficients and equations (2.7) through (2.10) to estimate ECS_{jt}. On average across the 250 replications, we obtain 1,790 estimates of ECS_{jt}.[12] We then merge the ECS_{jt} estimates with the hedonic data, creating 250 datasets with which to estimate equation (2.11). We combine the estimates from these regressions using Rubin's Rule (Rubin 1987), reporting the mean of each resulting vector of coefficient estimates in the results tables.

To calculate the standard errors, we first estimate the within imputation variance as $\text{Var}_{\text{within}} = \sum_{k=1}^{250} SE_k^2/250$ and between imputation variance as $\text{Var}_{\text{between}} = \sum_{k=1}^{250}(\beta_k - \bar{\beta})^2/(250-1)$, where k indexes the individual replication sample. We then calculate the total variance as $\text{Var}_{\text{total}} = \text{Var}_{\text{within}} + \text{Var}_{\text{between}} + \text{Var}_{\text{between}}/250$.The standard error is the square root of

[12]There are 2,083 zip-code-years in the data with some visitation, but our data-generating process excludes some zip-code-years from each replication dataset.

total variance. The individual replication standard error estimates are clustered by property or zip code, depending on the model.

The repeat sales model is not without its challenges. Only a subset of housing units in the data have sold more than once, given the limited market and time period of the study, and homes that sold more than once may have different attributes than properties that sold only once (or not at all) in given study period. Thus, restricting the sample to repeat sales may produce a selective implicit price (Freeman et al. 2014). The 300,207 repeat sales in our data, and 146,903 sales in the sample for our baseline model, account for more than 50% and more than 25% of qualified sales in our original data, respectively. Thus, they may be reasonably representative of the housing market in the Tampa metropolitan area. Figure 2.3 shows that average repeat-sales prices (red line) and average prices for all observed property transactions (blue line) in the three counties during the study period have very similar trends over time.

2.5.2.2 Long-Difference Model

Although the property fixed-effects model represents an advance over the cross-sectional approaches in the literature on water quality, we make an additional modification in the interest of better matching the econometric approach with the theoretical model in Section 3. Recall that in that model, households make short-run recreation decisions, conditional on home location, based on water quality at recreation sites. Thus, the impacts of recreational water quality on visitation are identified from short-run variation. The home purchase decision, however, is a long-run choice, and home prices should capitalize the expected long-run benefits from recreation and amenity value of water quality. In our view, a long-difference hedonic model may better fit this theory than the standard fixed-effects approach described above, in which the identifying variation for both the local water quality and recreational utility parameters is short-run.

The climate literature uses long-difference models to identify the impacts of medium- to long-run variation in temperature on economic outcomes of interest (Dell et al. 2014). Following this literature, we construct average long-run housing price changes, average local water quality changes, and average recreational utility index changes in two time periods for the same property. The two time periods in our main long-difference specification are 1998–2003 and 2009–2014, which we refer to as period a and period b, respectively. We choose these two periods because they represent the earliest and latest six-year periods in our data, and because doing so allows us to avoid the unusual property price changes during the housing boom and bust, evident in Figure 2.3. Our long-difference sample is very small because inclusion in the sample requires that a property sell at least once in *both* the early and late time periods, so that we can take differences, instead of at least twice in the full sample (the inclusion constraint in the property fixed-effects model). In addition, we only observe Manatee County transactions beginning in 2005, so properties in this county drop out entirely. In Figure 2.3, we can see that the average price in the long-difference sample (green lines) starts out somewhat lower than the full and repeat-sales samples in 1998, but then tracks very closely with the larger groups through 2003 (the pre-housing crisis, early period used for differencing). Post-housing-crisis, the price level for this small sub-sample is again slightly lower than the larger groups, but the trends over time are very similar.

Our approach estimates the average price of property i during time period a as:

$$\overline{P_{ija}} = \frac{1}{n}\sum_{t\in a} P_{ijt}, \tag{2.12}$$

where n is the number of times the property sold in time period a. We then construct the following equation:

$$ln\overline{P_{ija}} = \theta_0 + \theta_1\overline{Age_{ija}} + \theta_2 ln\overline{WQ_{ija}} + \theta_3\overline{ECS_{ja}} + \alpha_i + \epsilon_{ija}, \tag{2.13}$$

where $\overline{WQ_{ija}}$, $\overline{Age_{ija}}$, and $\overline{ECS_{ja}}$ measure the average local water quality, average property age and average recreational utility index of property i in zipcode j during period a. An analogous equation can be written for period b, in which the subscript a in equation (2.13) is replaced with the subscript b.

Differencing the two time periods drops the time-invariant property fixed effect α_i and results in:

$$\Delta lnP_{ij} = \theta_0 + \theta_1\Delta Age_{ij} + \theta_2\Delta lnWQ_{ij} + \theta_3\Delta\widehat{ECS_j} + \Delta\epsilon_{ij}, \tag{2.14}$$

where ΔlnP_{ij} is the change in the log housing price of property i in zipcode j between period a and period b. The independent variables are interpreted in a similar way.

The coefficients of interest are θ_2 and θ_3, which measure how long-run changes in local water quality and recreational opportunities affect the housing price. Similar to β_2 and β_3 in equation (2.11), we expect θ_2 and θ_3 to be positive. The interpretation of θ_2 and θ_3 is complicated by the fact that the dependent variable is the difference in log prices, and our independent variables are differences in log water quality and average ECS. We interpret the coefficients as marginal effects, instead of actual MWTP estimates. For instance, we interpret θ_2 as the marginal effect of the average water quality increase from period a to period b on the average housing price in period b, holding constant the average housing price and water quality in period a.

As we did for the property fixed-effects models, we use multiple imputation to obtain coefficient estimates and associated standard errors for equation (2.14), in order to address potential measurement error from estimating ECS_j at the zipcode, rather than the property level, following the procedure described in Section 5.2.1.

2.6 Results

2.6.1 Demonstration of the Typical Hedonic Approach

Before estimating our preferred two-stage model, we start with a demonstration of the typical hedonic approach, providing some analysis to support the heuristic critique we developed around Figure 2.1. We estimate equation (2.11), leaving out the recreational utility component (ECS_{jt}), and assigning water quality monitors to properties as long as they are within a specified radius of the home—defined at 1, 2, 3, 5, and 10 km—ignoring actual recreation behavior. We do this, first, using only the local water quality monitors, and leaving out the recreational water quality monitors in Tampa Bay. Next, we run the same set of regressions using all (local and recreational) water quality monitors within the specified radii, so that the contribution of recreational waters to homeowners' WTP for pollution abatement can be captured within the five different radii.

Coefficient estimates and their 95 percent confidence intervals, measured against the lower horizontal axis, are reported in Figure 2.4, with the local-monitor results in blue and the all-monitor results in red. The sample size for each regression is reported above each estimate; sample size grows with the specified radius for the "zone of influence" because there are many fewer properties with reporting water quality monitors a short distance away, so the number of property transactions with reporting monitors grows as we draw larger circles. To ease interpretation, the implied MWTP for the observed average 10% increase in DO in the Tampa Bay watershed from 1998-2014 can be read from the upper horizontal axis.

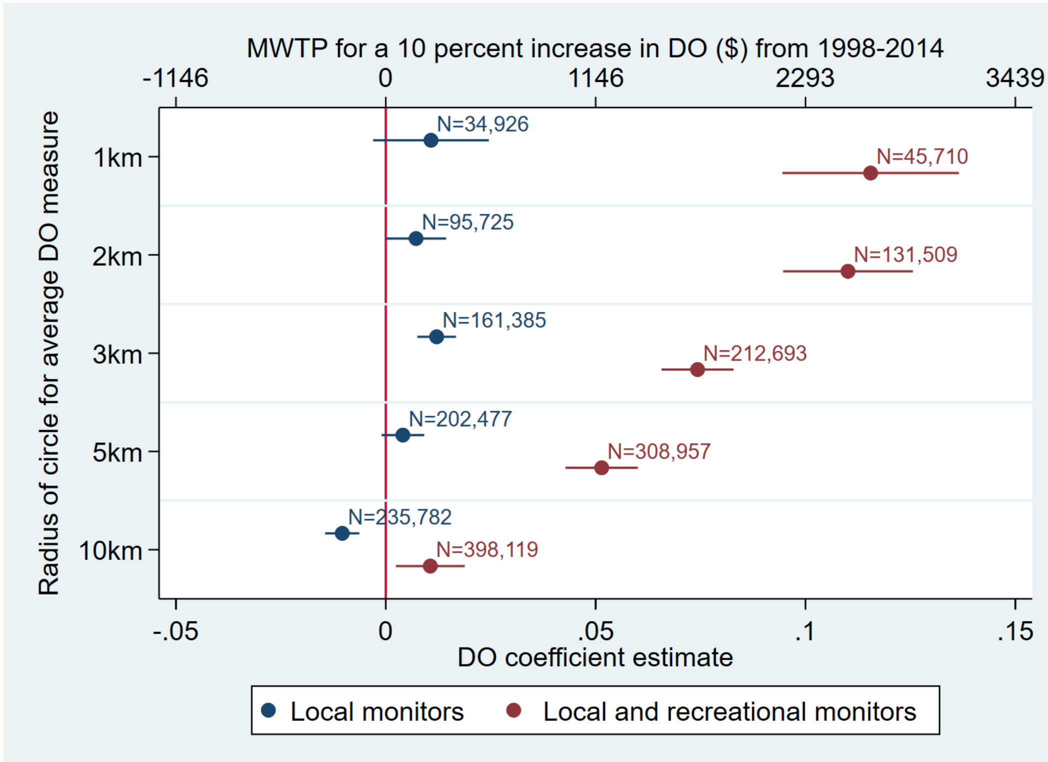


Figure 2.4: Coefficient estimates and MWTP for DO from a typical hedonic approach when water quality measurements are averaged for monitors within varying radii of properties.

Notes: Red point estimates and confidence intervals are from regressions using only local water quality monitors. Blue point estimates and confidence intervals are from regressions using all (local and Tampa Bay) water quality monitors. Sample sizes for each regression are reported above each estimate.

Several insights arise from Figure 2.4. First, estimated coefficients for the local measures of DO are mostly small and positive, hovering around 0.01, for an implied MWTP for the average water quality improvement in the watershed during the study period of $100-$300. Second, once we include the recreational monitors, the estimated coefficients increase appreciably. For the smallest radii, the red coefficient estimates in Figure 2.4 imply a MWTP for the average water quality improvement from 1998 to 2014 of between \$2,500 and \$3,000 per property—an order of magnitude larger than those without the recreational waters included. Third, the all-monitor coefficients decrease in magnitude as we draw larger and larger circles around Tampa Bay homes to describe water quality, until the estimates at 10 km are in the range of the local-monitor-only estimates. This is consistent with the issue we raised in the heuristic discussion of Figure 2.1; it may be the case that the larger "zones of influence" capture so many irrelevant water quality monitors (those that describe water quality in locations that a household does not value) that the signal of pollution abatement's value at key sites gets lost in the noise from the sites with little or no value.

This last insight is also consistent with households further from water having systematically lower MWTP for water quality improvements. As we allow monitors at increasing distances from each property to influence our coefficient estimates, we are also able to include homes in the sample that are further and further from any water quality monitor, and are thus further from water altogether. If heterogeneity in MWTP among property-owners depending on water proximity explains the observed pattern of decreasing "all-monitor" coefficient estimates, then it is a real phenomenon that one would want to capture in any estimate of the monetized benefits of water pollution abatement. We do not observe this pattern consistently in the local-monitor results, however. The value of local water pollution abatement is quite similar when we use monitors between 1 km and 10 km.

Though this analysis—comparable to the standard hedonic approach when valuing changes in pollution—points to the importance of recreational benefits in estimating MWTP for water pollution abatement, it is naïve relative to actual recreation behavior. A different approach is needed to match homes with the recreation sites that individuals living in those homes typically visit. Thus, we implement the two-stage approach described in Sections 3 and 5.

2.6.2 Main Model: First Stage Recreation Demand Results

Results from the recreation demand model are reported in Table 2.2. Columns 1 and 2 define water quality as the average DO at Tampa Bay monitors within 3 km of each fishing site in a given year; column 3 uses a 5-km radius to define average water quality in a site-year. Column 2 reports estimates from a model using travel time, rather than travel cost, as the relevant travel cost variable, an alternative approach in the literature (Cesario 1976, Wilman 1980), using the 3-km radius. Results are robust to these differences in specification. As the travel cost to a site increases by $1, the probability of an angler fishing at the site decreases by about 11%. In columns 1 and 3, this coefficient can be interpreted as the marginal utility of income, and our estimate is similar to others in the literature (von Haefen 2003). In the recreation demand model, there is little difference in the estimates when we use monitors within 3km of a fishing site to describe water quality, or those within 5km.

Table 2.2: First-stage recreation demand model

	(1) Travel cost (3km)	(2) Travel time (3km)	(3) Travel cost (5km)
Travel cost (US dollars)	-0.110***		-0.113***
	(0.00078)		(0.00080)
Travel time (minutes)		-0.0640***	
		(0.00044)	
DO (mg/L)	0.0722***	0.0789***	0.0663***
	(0.0089)	(0.0083)	(0.0114)
Seagrass abundance	-0.170***	-0.135***	-0.141***
	(0.0104)	(0.0103)	(0.0104)
Alternative-specific constants	Yes	Yes	Yes
Observations	1,765,796	1,765,796	1,801,615

Standard errors in parentheses.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: Models are estimated using conditional logit, with a choice set of 85 fishing sites visited during the study period. Columns 1 and 2 link fishing sites to Tampa Bay water quality monitors within 3km. Column 3 links sites with monitors within 5km. Travel cost is estimated as the sum of the value of travel time (1/3 of foregone wages times round-trip travel time) and the operational cost of travel (AAA's driving cost times round-trip distance).

The effect of DO on visitation is positive, significant, and very similar across all three models. Anglers from the three counties are 6.6 to 7.9% more likely to recreate at a site if the DO level increases by 1 mg/L, equivalent to a 16% increase from mean DO at the Bay monitors in our sample.

The coefficient on seagrass abundance is statistically significant and negative. A 1-unit increase in seagrass abundance (a 43% increase over the mean) lowers the probability of fishing at a site by 13.5 to 17%. Though seagrass abundance is correlated with high water quality in Tampa Bay, the negative coefficient may be due to the fact that seagrass can be a disamenity to anglers. For example, anglers in shallow water must take care not to scar seagrass beds with a boat motor's propeller, and seagrass can catch and tangle fishing lines.[13]

In the online Appendix, we re-estimate the recreation demand model using the 5 mg/L DO threshold instead of the continuous DO concentration (Table A.3). Results are similar, except that the local DO coefficient is not significantly different from zero in the model using a 5-km radius to describe water quality at fishing sites.

Using the estimated parameters in Table 2.2, we then estimate the expected utility from recreation trips initiating from each zip code j in each year following Equation (2.8) through Equation (2.10). The average value of expected utility (ECS_{jt}) from the RUM model calculated from trips occurring in each zip code-year is \$35.98 (Table 2.1).

Figure 2.5 maps the mean values over the study period of our recreational utility estimate by zip code, as well as mean DO values at each recreational fishing site. The heat map of ECS_{jt} quintile by zip code shows some predictable results. For example, values are high in Pinellas County (the peninsula that separates the Bay from the Gulf). Some other coastal zip codes, especially those in southern Hillsborough County, also have high average recreational utility. Figure 2.5 also shows, however, that residents of the region's less densely-populated zip codes further from the coast (e.g., in north-

[13] Damaging seagrass beds in Florida can result in a fine of up to \$1,000 (see: https://www.flseagrant.org/news/2016/07/savanna-barry-smart-boating-seagrasses-important/).

ern Hillsborough and eastern Manatee County) also obtain high utility from recreational fishing in the Bay. This is not surprising, given the relatively high average travel time (about 90 minutes round trip) to fishing sites in the MRIP sample. It does support our contention, however, that accounting for actual recreation behavior may paint a different picture of the value of recreational water quality than approaches that proxy for behavior using proximity.

To estimate the marginal effect of DO increases at Tampa Bay recreational fishing sites using the RUM model, we can recalculate the ECS_{jt} using the DO coefficient estimate from column 1 of Table 2.2 and use equations (2.7) through (2.10). Given that the mean DO level in the Tampa Bay watershed increases by about 10% from 1998-2014, we use Equation (2.9) to estimate the change in ECS_{jt} associated with this increase in water quality. The increase in ECS_{jt} is \$0.42 per trip on average, which is about a 1.2% increase over the mean in the expected utility of recreation.

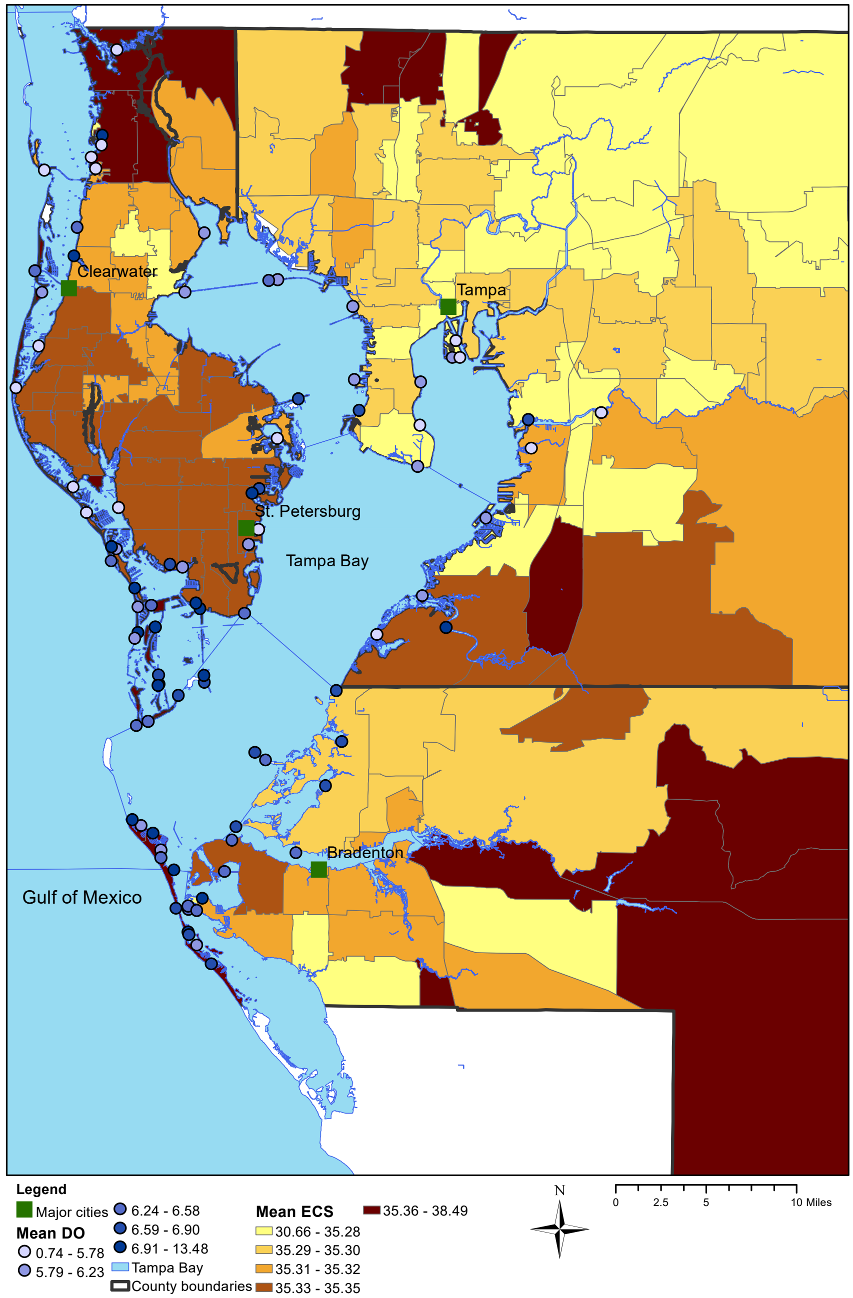


Figure 2.5: Average ECS and average DO (mg/L), 1998-2014

Notes: This figure maps average ECS and average DO for all zip codes and recreational fishing sites, respectively, associated with properties in the repeat-sales sample.

2.6.3 Main Model: Second-stage Hedonic Results

Results from estimating Equation (2.11) are reported in Table 2.3. In column 1, our baseline property FE model, a 10 percent increase in local DO leads to a 0.114% increase in mean property prices.[14] Households' MWTP for local DO is about $261 per property for the observed 10% increase in DO from 1998-2014. This is in line with the previous literature's small, positive estimates of MWTP for local water quality improvements.[15]

The recreational utility index coefficient in column 1 of Table 2.3 is large and statistically significant. From the previous section, a 10% increase in DO is associated with a $0.42 increase in ECS_{jt}. From column 1 of Table 2.3, a $1 increase in ECS_{jt} is associated with a 24.9% increase in the housing price, 1998-2014. Thus, the $0.42 increase in ECS_{jt} is associated with a 10.46% increase in the average housing price, or about $23,980 per property – almost two orders of magnitude larger than our estimated MWTP for local water quality improvements.

One test of whether our amenity and recreational estimates are really separable is to observe what happens to our estimates of the amenity value of local water quality improvements when the recreational utility index is omit-

[14]Results using the 5 mg/L DO dummy at a 3-km radius are similarly positive and significant (online Appendix, Table A.5).

[15]If we estimate the typical hedonic model using a vector of property characteristics instead of property fixed effects, we obtain intuitive results for property characteristics, and counter-intuitive results for water quality—better water quality has a negative, insignificant effect on property values. Results are reported in the online appendix, Table A.4. This underscores the importance of controlling comprehensively for property characteristics when estimating MWTP for water pollution abatement.

ted from the model. Column 2 of Table 2.3 shows that the local DO coefficient is insensitive to the exclusion of ECS_{jt}. This suggests that the two parameters are, in fact, picking up different aspects of MWTP for water quality improvements.

In column 3 we repeat the model from column 1, except that we use a 5-km radius to characterize water quality at local waterbodies and Bay recreational fishing sites. The sample size grows for this model, because we can now include repeat-sales properties that are located between 3 and 5km from at least one local water quality monitor. In column 3, neither the local DO nor the ECS_{jt} coefficient are statistically different from zero, and both are smaller than their counterparts in column 1.

In column 4 we report estimates that use the travel-time parameters from the RUM model (column 2 of Table 2.2) to construct the recreational utility index, rather than using the travel cost estimates (column 1 of Table 2.2). The coefficient on local DO does not change, while the ECS_{jt} estimate is less than one-half that in column 1. The coefficient on ECS_{jt}, together with the DO coefficient estimate in column 2 of Table 2.2, implies a MWTP for recreational water pollution abatement of about \$10,200 per property.

In the online Appendix (Table A.5), we re-estimate the property FE models in Table 2.3, using the 5 mg/L DO threshold. Results are qualitatively similar and perhaps a bit stronger. The coefficient on local DO is small, positive and significant using a 3-km radius, and positive and weakly significant at 5km. The coefficient on ECS_{jt} is large, positive and significant at 3km, and also at 5km.

Table 2.3: Second-stage hedonic regression results

	(1) Basic 3km	(2) No ECS_{jt} 3km	(3) Basic 5km	(4) Travel time for ECS_{jt}	(5) County time trend	(6) Subdiv. time trend
ln(DO)	0.0114***	0.0111***	0.00155	0.0116***	0.0109***	0.00956***
	(0.00307)	(0.00307)	(0.00416)	(0.00312)	(0.00320)	(0.00341)
ECS_{jt}	0.249***		0.133	0.106**	0.0617	0.0280
	(0.0814)		(0.0824)	(0.0450)	(0.0796)	(0.0847)
Property age	-0.0123***	-0.0123***	-0.0105***	-0.0122***	-0.0126***	-0.0139***
	(0.00328)	(0.00328)	(0.00314)	(0.00334)	(0.00327)	(0.00336)
Property FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-year trend	No	No	No	No	Yes	No
Subdivision-year trend	No	No	No	No	No	Yes
N	146,903	146,903	183,582	146,903	146,903	126,926
R-squared	0.626	0.626	0.627	0.626	0.632	0.631
MWTP for 1 mg/L local DO ($)	451	440	62	465	432	379
MWTP for 1 mg/L Tampa Bay DO ($)	37,468	N/A	20,013	15,950	9,284	4,213

Estimated standard errors in parentheses are clustered by property.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: The dependent variable is the log property transaction price. Column 1 uses a 3-km radius to define average water quality around properties. Column 2 drops the recreational utility index, ECS_{jt}. Column 3 repeats column 1, using a 5-km instead of a 3-km radius to define average water quality around properties. N rises in column 3 because more repeat sales are located within 5 km of at least one water quality monitor than within 3 km. Column 4 uses travel time instead of travel cost in the first stage to estimate ECS_{jt}. Column 5 includes county*year trends as additional controls. Column 6 includes census subdivision*year trends. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in detail in Section 5.2.1.

2.6.3.1 Models with interactions between time trends and spatial controls

One threat to identification in our baseline model is that we do not control for factors influencing housing prices that vary over both time and space that could be correlated with water quality. For example, the economic recession and housing crisis that occurred during our sample period and are visible in Figure 2.3 may have had heterogeneous effects by county or neighborhood, and those effects could be correlated with water quality and recreational utility. This would bias our coefficient estimates. The most comprehensive approach to this challenge would interact our year fixed effects with geographic controls at a higher spatial scale than the property. However, given that recreational utility is estimated at the zip code level, and there are only a small number of zip codes in the data (and even fewer counties, subdivisions, or other levels of spatial aggregation), this approach leaves too few repeat sales to identify the effect of recreational utility on property prices.[16] We estimate two alternative models that include different levels of interactions between a time trend and geographic controls, as well as year fixed effects, as in equation (15):

$$lnP_{ijct} = \beta_0 + \beta_1 Age_{it} + \beta_2 lnWQ_{it} + \beta_3 \widehat{ECS}_{jt} + \alpha_i + \gamma_t + \lambda_c * T + \epsilon_{ijct} \quad (2.15)$$

[16]There are 138 zip codes in the data (66 in Hillsborough County, 26 in Manatee, and 46 in Pinellas). There are 16 census subdivisions in the three counties: 7 in Hillsborough, 4 in Manatee, and 5 in Pinellas.

where λ_c indicates that a property is located in county or census subdivision c, and T is a linear time trend. Results are presented in Table 2.3, columns 5 and 6.

The effect of local DO on property prices, identified using variation in water quality over time within 3 km of a property, and within a year, is relatively insensitive to the inclusion of these additional controls. The estimated effect of recreational utility on property prices changes substantially, however. The coefficient on ECS_{jt} in column 5 is about one-fourth the magnitude of that in our baseline model in column 1. In column 6, when we include the subdivision-specific trends, it is even smaller and also statistically insignificant.

The identifying variation for the recreational utility index coefficient comes from changes in recreation within zip codes over time. In column 5, the trend in recreational utility (which may be more important to property owners than periodic departures from the trend) at the county level is removed from the identifying variation for the recreational utility index. The model in column 6 is even more restrictive. In our view, the variation in recreational utility across census subdivisions and counties is important variation to capture in the coefficient on ECS_{jt}, thus the column 1 model is preferred to those with the county-specific trend and subdivision-specific trend controls. However, concerns about identification in that model provide additional support for our long-difference approach.

2.6.4 Long-difference models

Table 2.4 reports results from the long-difference models. Column 1 reports results from our baseline long-difference model, with standard errors clustered by property. In the baseline model, if we hold average property price and average local DO in the first time period (1998-2003) constant, a 1% increase in average local DO in the second time period (2009-2014) is associated with a 0.0226% increase in the average second-period property price. If the average second-period local DO increases by 10% (the average change observed in the data between period a and period b), the average property price increases by 0.226%. From Table 2.1, the average property price from 2009-2014 is about \$196,000. Thus the marginal effect of an 10% increase in average local DO from the first to the second period is about \$440 per property, almost a 70 percent increase over the \$261 estimate using column 1 of Table 2.3. Note that the property FE model estimates MWTP for a change in DO using short-run variation, while the long-difference models in Table 2.4 exploit the change between period a and period b, with more than a decade separating the midpoints of the two periods.

Interpreting the ECS_j coefficient in column 1, recall that a 10% increase in DO is associated with a \$0.42 increase in ECS_{jt}. From Column 1, a \$1 increase in ECS_j is associated with a 1.19% increase in property prices. Thus, the \$0.42 increase in ECS_{jt} is associated with a 0.5% increase in the average housing price, or about \$1,000 per property. Property markets appear to have capitalized a value of regional recreational fishing benefits from water quality improvements in Tampa over two decades that is more than twice the size of the value of improvements in ambient water quality very near to properties. Again, the property FE ECS_{jt} coefficient and the long-difference coefficient measure different things. But if we compare the two qualitatively, the recreational utility component of the estimated value of water quality improvements in the long-difference models is much smaller than in the property FE models.

In column 2, we estimate the same long-difference model (equation (2.14)), clustering standard errors by zip code, rather than by property. While we measure property-specific local DO, the recreation demand index is constructed by zip code, so it may be more appropriate to cluster at this higher level of aggregation (Cameron & Miller 2015). Our estimate of the marginal effect of local DO is no longer significant, but the ECS_{jt} coefficient is.

As we did for the property FE models, we also try dropping ECS_j from the long-difference model (in column 3). The effect of improved local DO on property prices is slightly smaller, but not different enough to raise a concern that the two variables are not identifying different things.

Table 2.4: Second-stage long difference models and hedonic regression results

	(1) Long diff. SE property	(2) Long diff. SE zipcode	(3) Long diff. No ECS_{jt}	(4) County time period	(5) Subdiv. time period	(6) Hedonic SE property	(7) Hedonic SE zipcode
$\Delta ln(DO)$	0.0226***	0.0226	0.0182***	0.0227***	0.0360***		
	(0.00645)	(0.0210)	(0.00642)	(0.00647)	(0.00772)		
ΔECS_j	0.0119***	0.0119***		0.0118***	0.0158***		
	(0.00130)	(0.00330)		(0.00130)	(0.00141)		
Δ Property age	0.0351***	0.0351***	0.0351***	0.0352***	0.0384***		
	(0.00178)	(0.00316)	(0.00174)	(0.00178)	(0.00190)		
ln(DO)						-0.000654	-0.000654
						(0.00568)	(0.0177)
ECS_{jt}						0.226	0.226
						(0.142)	(0.261)
Property age						-0.0127	-0.0127
						(0.0142)	(0.0270)
Property FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	No	Yes	Yes
County-time period dummy	No	No	No	Yes	No	No	No
Subdivision-time period dummy	No	No	No	No	Yes	No	No
Cluster SE level	Property	Zipcode	Property	Property	Property	Sales	Zipcode
N	14,390	14,390	14,390	14,390	12,219	32,024	32,024
R-squared	0.049	0.049	0.041	0.049	0.062	0.206	0.206
MWTP for 1 mg/L local DO ($)	751	751	605	755	1,197	0	0
MWTP for 1 mg/L Tampa Bay DO ($)	1,529	1,529	N/A	1,516	2,030	34,007	34,007

Standard errors in parentheses are clustered by property unless otherwise noted.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: Columns 1-5 report results from long-difference models, using a 3-km radius to characterize water quality and clustering standard errors by property. Column 2 clusters standard errors by zipcode. Column 3 drops the recreational utility index, ECS_{jt}. Column 4 includes differences in county*time period (a and b) as additional controls. Column 5 includes census subdivision*time period controls. Columns 6 and column 7 show results from property fixed-effects models estimated using the long-difference sample, and clustering standard errors by property and zipcode, respectively. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in detail in Section 5.2.1.

In the next two columns of Table 2.4, we include a set of interactions between counties and the two time periods (column 4) and a set of interactions between census subdivisions and the two time periods (column 5). Our reasoning is similar to that in equation (15); if inter-temporal shocks common to counties or neighborhoods are correlated with water quality improvements from period a to period b, this could bias our estimates of the effects of water quality on property prices. The inclusion of these additional covariates does not change our estimates much compared to column 1. If anything, the estimates increase a bit in the model with subdivision*period controls (column 5).

In the final two columns of Table 2.4, we re-estimate the basic property FE model from Table 2.3, restricting the sample to the same properties that appear in the long-difference sample, and clustering at two levels (property and zip code). (These models still use more observations than the long-difference models, because we include all transactions from 1998-2014 for these homes, and not just those that occur in the first and last six years of the sample period). We cannot directly compare the coefficient estimates in columns 6-7 with those in columns 1-2. However, we note that when using the long-difference sample with the property FE approach, neither coefficient is statistically different from zero. This suggests that the differences between our long-difference estimates and our FE estimates may be due both to different samples, and to different specifications.

In the Appendix, Table A.6, we also test the robustness of the long-difference results to different definitions of period a and period b: allowing the period a sample to extend to just before the recession and housing crisis (period a: 1998-2006 and period b: 2009-2014), and splitting the full time period in half (so: 1998-2007 and 2008-2014). In both cases, the marginal effects of both local DO and recreational utility are quite a bit larger than our estimates in Table 2.4. The sample sizes are also about twice those in Table 2.4. As in column 2 of Table 4, When we cluster standard errors by zip code in these models, the marginal effect of local DO (a small share of our total estimated value of water quality improvement) is not statistically significant. Both of these alternative models include transactions during the housing boom, and the second approach also includes the subsequent bust. Thus, we report the Table 2.4 results as the main long-difference results.

Given its comportment with the theoretical model and ability to control comprehensively for unobservables, the long-difference model may be the preferred approach to valuing water quality improvements. However, the choice between long differences and property FEs creates a stark tradeoff in sample size (and possible selection). The long-difference repeat-sales sample is less than one-tenth of the full sample, because we must observe properties sold at least once in each time period, as well as recreation at fishing sites from each zip code in each time period, in order for those properties and zip codes to be included in the models. Figure 2.3 suggests that the trends in property prices in the long-difference sample are similar to those for all sales and repeat sales in both periods, with the exception of a low start in 1998 for the long-difference sample. However, Figure 2.6 shows that the long-difference sample drops many zip codes entirely, including all zip codes in Manatee County, many of which have high average recreational utility from fishing in the Bay (mapped in Figure 2.5). As noted earlier, Manatee County drops out because we only observe transactions there from 2005-2014. Figure 2.5 also shows that average water quality at coastal fishing sites in Manatee County are among the highest in our sample. Thus, the long-difference sample may not be representative of Tampa area property owners' willingness to pay for water quality improvements, especially in recreational waters. The exclusion of these properties likely matters for the magnitude of our estimates, and may explain some of the differences between the property FE and long-difference results. For these reasons, in the benefit-cost analysis in Section 7, we use both the property FE and the long-difference results, reporting a range of estimates.

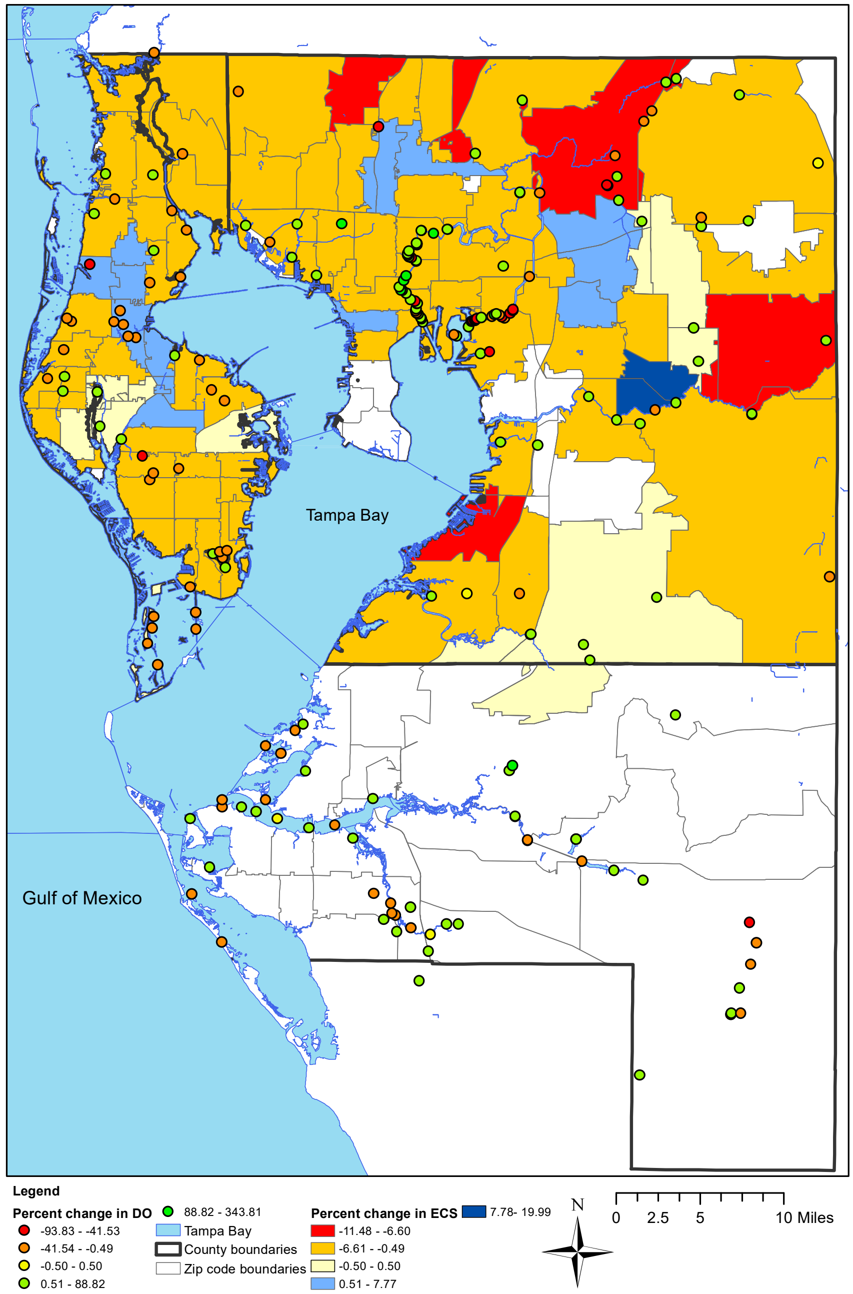


Figure 2.6: Percent change in average ECS and percent change in average DO, 1998-2003 to 2009-2014

Notes: This figure maps the change in average ECS and the change in local ambient DO for all zip codes associated with properties in the long-difference repeat-sales sample.

2.6.5 Robustness Checks

2.6.5.1 Effects of proximity to water

Although Equations (2.11) and (2.14) capture the overall effect of water quality on property prices, they do not allow us to examine how this effect varies with proximity to water, as is common in the prior literature. To this end, we also estimate models that interact both local DO and the recreational utility index with a property's distance to water. Table A.7 in the online Appendix reports results. The FE model is in column 1, and the long-difference model is in column 2. In both columns, higher DO in local water raises property prices, and the effect of recreational utility is also positive and significant. The FE model suggests that recreational value falls with distance from water, but amenity value does not. The long-difference model indicates no effect of distance on the marginal value of water quality improvements.

2.6.5.2 Allowing for recreation in local waterbodies

One potential challenge to our results is that some local waterbodies may also provide recreation opportunities. This would not be a problem for our estimates in a benefit-cost analysis—a comprehensive estimate is desirable. But to claim that our local MWTP estimate is primarily an amenity value while the recreational index captures recreational value, further analysis is needed. Thus, we divide local waterbodies within close radii of properties into those in or near public parks, and those well outside of parks. To categorize local park monitors, we use the Florida parks shapefile from OpenStreetMap, an online map database built and maintained by volunteers worldwide (OpenStreetMap contributors 2018). We define park monitors as monitors within 150 m of park boundaries. We estimate the mean annual DO concentration of waters in all parks within 3km of a property, and join this new time-varying water quality measure with property sales. We then include the log-transformed local park water quality measure in the property FE model. Note that we do not observe recreation at these sites, so we are stuck using the standard, naive approach to estimating the benefits of local park aquatic recreation by proximity.

Table A.8 in the online Appendix reports the results of a property FE model and a long difference model including the mean park DO concentration. Because DO is not measured in many neighborhood ponds, our sample size shrinks to fewer than 42,000 repeat sales, and fewer than 3,600 long-difference repeat sales. Note that in this selected sample of homes near parks with water quality data, the results may say more about homeowners who locate near parks than it does about average MWTP for water quality. It is difficult to conclude much from this test, given the very different results in the two models, and the very small and select sample. Nonetheless, if anything, including local (non-Bay) recreational waters in the analysis increases the coefficient estimate on the Bay recreational utility component of our models. Given the long-difference results in Table A.8, it is possible that our local DO estimates in the main models are picking up some recreation benefits, for homes near parks with non-Bay aquatic recreation.

2.6.5.3 Smaller spatial radii for local water quality monitors

Recall that our main models take the average of all monitors within a 3-km radius of a property to represent local water quality, with additional results reported for a 5-km radius. To further test the robustness of our results to this choice, we estimate models using radii of 1 km, 500 m and 300 m. Table A.9 in the online appendix reports results for both continuous DO and the 5mg/L threshold. We report only the local DO results, with the full set of coefficients available on request. In the property FE models, the property value effect of local water quality gets larger as the radius gets smaller, to 500m, consistent with previous literature indicating larger effects for properties closer to the water (Walsh et al. 2011, 2017, Wolf & Klaiber 2017). The effects lose significance for the smallest radius, likely due to the very small number of observed repeat sales within 300 m of one or more reporting water quality monitors. In the long-difference models, the impacts of local DO are not statistically significant below a 3-km radius—these samples of homes sold at least once in each period that are located very close to monitors are very small.

2.6.5.4 Moving average DO concentrations

While it is common in the literature to use the average water quality measure from the calendar year of a property's sale to represent water quality conditions in hedonic regressions, as we do above, we implement a set of robustness checks using average DO concentrations within a 3 km radius of a property 3 months, 6 months, and 1 year before each sale date, reporting results in Table A.10 in the online appendix.[17] Column 1 uses the mean DO concentration in the 12 months prior to a property's sale date. The magnitude of the coefficient on $ln(DO)$ in this model is very similar to that in our baseline model in column 1 of Table 2.3. Results from column 2 of Table A.10 suggest statistically insignificant MWTP for local DO increases 6 months prior to a sale. Column 3 indicates (counter-intuitively) that local DO improvements 3 months prior to a sale are actually associated with lower property prices. Property transactions can take several months, so homeowners may have no or low MWTP for local DO increases while waiting to close transactions. However, we would not expect water pollution abatement to reduce property values in this shortest window. The coefficients on ECS_{jt} in Table A.10 show consistently that homeowners have positive and statistically significant values for recreational utility (a function of Bay water quality), and that these values are robust to varying time windows prior to a sale, though somewhat smaller than in our baseline model.

[17]The number of observations varies by specification in Table A.10 because the number of repeat sales for which we are able to estimate average water quality in each time window varies. Note that the 12-month moving DO average (column 1) actually gives us a larger sample than the models in Table 2.3 (which use average DO in the calendar year of each property transaction). Also, in Table A.10, as the time window for calculating the DO average shrinks, so does the sample. This is due to the fact that some monitors do not report very frequently, and properties are dropped from the sample when there are no DO observations in the relevant time window.

2.7 Discussion and Conclusions

Our empirical results demonstrate that valuation of water pollution abatement using hedonic analysis is strongly downward-biased if recreational waters are omitted, and if they are included but the analysis ignores actual recreation behavior. Taken together, our integrated two-stage model and robustness checks suggest that increases in dissolved oxygen (DO) improve both recreational and aesthetic amenities and that homeowners in Tampa Bay have significant MWTP for both of these improvements. Our baseline MWTP estimates for recreational water quality improvements in Tampa Bay from 1998-2014 are much larger than our estimates of MWTP for local amenity improvements. The two effects appear to be separable in Tampa Bay, suggesting prior hedonic studies of the value of water quality could provide unbiased estimates of local amenity values, but may exclude the potentially much larger regional recreational values.

From 1998-2014, the average DO concentration in the Tampa Bay region increased by about 10%. Table 2.5 summarizes our monetized benefit estimates for this water quality improvement. Using the property FE estimates reported in column 1 of Table 2.3 (panel A), the local amenity benefits from this improvement range from about $16.8 million if we apply them only to our sample households, to about $210.4 million if we apply them to all owner-occupied households in the Tampa Bay metro area in the 2010 Census (U.S. Department of Housing and Urban Development 2015). In contrast, when we use the two-stage model coefficient estimates to also monetize the recreation benefits from DO improvements over the 16-year study period, our benefit estimates range from \$1.5 to \$19.3 billion, depending on the scope of the property market to which these benefits accrue.

In panel B of Table 2.5 we use the long-difference estimates from Table 2.4 for the same exercise. The monetized estimates of MWTP for a 10% local DO improvement range from \$6.4 to \$356.3 million, depending on the geographic scope of homeowners to whom benefits accrue. The MWTP estimates for recreational improvements are substantially smaller than those calculated from our baseline estimates (from \$20.4 million considering only our long-difference sample properties to \$1.2 billion for the whole Tampa metro area). At the low end, they are smaller due to both a much smaller repeat-sales sample, and smaller coefficient estimates. At the high end, the difference is due only to our differences in estimates.

How do these benefit estimates compare to the costs that firms, homeowners, governments (taxpayers), and other stakeholders incurred to achieve the water quality gains observed in the Tampa Bay watershed between 1998 and 2014? There are several challenges to answering this question. First, the water pollution control projects that contributed to DO gains over this period were incredibly diverse in scope and type, implemented by dozens of different public and private sector institutions (Beck et al. 2019). Second, the impact of each project has not been rigorously evaluated to determine its causal impact on water quality. One working paper suggests that some types of projects, particularly point-source nitrogen controls, may be statistically associated with

Table 2.5: Range of monetized benefit estimates for the observed 10% increase in average DO concentration in the Tampa Bay watershed, 1998-2014

Panel A: Benefit estimates using coefficient estimates from property fixed-effects model

Water quality benefit	**WTP – sample households only**	**WTP – all repeat sales, 1998-2014**	**WTP –Tampa Bay metro area**
Amenity benefit	\$261 x 64,353 = \$16.8 million	\$261 x 170,192 = \$44.4 million	\$261 x 806,000 = \$210.4 million
Recreational benefit	\$23,980 x 64,353 = \$1.54 billion	\$23,980 x 170,192 = \$4.08 billion	\$23,980 x 806,000 = \$19.3 billion
Amenity + recreational benefit	\$24,241 x 64,353 = \$1.56 billion	\$24,241 x 170,192 = \$4.13 billion	\$24,241 x 806,000 = \$19.5 billion

Panel B: Benefit estimates using coefficient estimates from long difference model

Water quality benefit	**WTP – sample households only**	**WTP – all repeat sales, 1998-2014**	**WTP –Tampa Bay metro area**
Amenity benefit	\$442 x 14,390 = \$6.36 million	\$442 x 170,192 = \$75.2 million	\$442 x 806,000 = \$356.3 million
Recreational benefit	\$979 x 14,390 = \$14.1 million	\$979 x 170,192 = \$167 million	\$979 x 806,000 = \$789 million
Amenity + recreational benefit	\$1,421 x 14,390 = \$20.4 million	\$1,421 x 170,192 = \$242 million	\$1,421 x 806,000 = \$1.15 billion

Notes: Sample households in panel A column 1 are the 64,353 homes sold twice or more in Hillsborough, Pinellas and Manatee counties from 1998-2014, and for which we also observe water quality variables and estimate the zip-code-level annual recreational utility index. Sample households in panel B column 1 are the 14,390 homes sold at least once between 1998-2003 and at least once between 2009-2014, and for which we also observe water quality variables and estimate the zip-code-level recreational utility index. All repeat sales in column 2 include all 170,192 homes sold twice or more in the three counties, 1998-2014. This includes properties dropped from our sample due to missing water quality or recreational visitation data. All homeowners in the Tampa Bay metro area in column 3 include the approximately 806,000 owner-occupied households in the 2010 Census in the Tampa-St. Petersburg-Clearwater metro area (U.S. Department of Housing and Urban Development 2015), which includes Hillsborough and Pinellas counties, as well as two counties excluded from our sample (Hernando and Pasco), and excludes Manatee County (which is in our sample). The amenity and recreational benefit estimates are calculated using the estimated coefficients from models reported in column 1 of Table 2.3 for panel A, and those reported in column 1 of Table 2.4 for panel B.

subsequent water quality improvements at downstream water quality monitors over time (Beck et al. 2019). Other approaches (for example, nonpoint source control and habitat restoration projects) may be less strongly associated with water quality improvements (Beck et al. 2019). However, we are not able to determine which projects actually caused the water quality improvements we observe in the data, and for which we estimate Tampa property owners' MWTP.

From the Tampa Bay Estuary Program, we obtained a catalog of the more than 800 projects implemented between 1971 and 2017. If we consider only those 600 projects implemented between 1998 and 2014 (our study period), and only those for which cost estimates exist (311 projects), the costs of these projects sum to about \$585 million, about 8% of our estimated benefits for all repeat-sales properties between 1998 and 2014, using the property FE results in panel A of Table 2.5. Using the long-difference results in panel B, estimated benefits are about 24% lower than this very rough cost estimate. If the benefits accrued more broadly—to all owner-occupied single-family homes in the metro area—then the benefits are about twice the costs, even using the long-difference results, which as noted earlier drop transactions for many properties with high average recreational utility. These are very favorable benefit-cost ratios when compared to other water quality benefit-cost analyses in the literature (Keiser & Shapiro 2019*b*,*a*, Keiser et al. 2019).

Though our benefit estimates are more comprehensive than prior work using hedonics or recreation demand modeling, they are still incomplete. We exclude the recreational fishing benefits that improved water quality has afforded non-residents such as tourists visiting Tampa Bay, as we used only the

MRIP survey data for anglers whose trips originated from zip codes in the three sample counties. Mitigating eutrophication also reduces emissions of methane, a greenhouse gas (Beaulieu et al. 2019), which could be valued using estimates of the social costs of GHGs and would almost certainly not be capitalized into local property prices. Moreover, the rebound in seagrass coverage in Tampa Bay results in additional nitrogen removal (as these healthy plants absorb nutrients for growth), generating a positive feedback. Scientists have estimated that the additional nitrogen removal services associated with the rebound in seagrass in Tampa Bay between 1982 and 2010 has, itself, removed enough nitrogen from the Bay to avert more than $20 million per year in expenditures for additional denitrification by municipal wastewater treatment plants and other sources (Russsell & Greening 2019). These kinds of avoided costs are also unlikely to be capitalized into local housing prices, as it would be difficult for homeowners to be aware of them. Thus, our benefit estimates are almost surely conservative.

We cannot assess the quality of the 311 available nutrient project cost estimates from the TBEP, or the projects for which costs have not been estimated. In addition, projects implemented before 1998 may contribute to the water quality changes we observe after 1998. Thus, costs may be over- or under-estimated.

This work adds to our understanding of how people value water quality improvements, especially nutrient pollution abatement. Eutrophication, a consequence of nutrient pollution, may cause large economic damages in the United States and elsewhere. Many local, state and federal regulations have been implemented to address this problem. Further work to help policymakers better understand how people value nutrient pollution abatement, and how these values are capitalized in housing markets, can contribute to a more comprehensive evaluation of such regulations.

We also contribute to the literature on hedonic valuation of pollution control, more generally. We estimate the first hedonic model valuing water quality that controls comprehensively and flexibly for property characteristics, using two different approaches. Our long-difference hedonics approach may comport better with hedonic theory than other approaches in the literature, given that the hedonic model considers the property location decision in long-run equilibrium. Lacking data on recreation site visitation at the property level—likely a problem faced by other researchers examining similar questions, unless they implement a household survey—we use multiple imputation to address the resulting measurement error relative to recreation data observed by property. These innovations may enable future work valuing water pollution and pollution control with a broader geographic scope than we have examined in this paper.

Chapter 3

The impacts of drinking water lead exposure on short-run and long-run educational outcomes

3.1 Introduction

The effects of environmental pollution on human health and welfare are fundamental parameters for designing and evaluating environmental regulation. The literature in epidemiology and economics documents significant negative impacts of contemporaneous air pollution on population health and welfare (see Currie et al. (2014) for a recent review). Studies suggest that early exposure to pollutants impacts long-term outcomes(Isen et al. 2017, Zaveri et al. 2019). However, only a few existing papers focus on the impact of water pollution on health and the long-term consequences of early exposure to drinking water pollution. It is important to understand these questions since existing studies suggest that drinking water quality is an ongoing problem, even in a developed country like the United States (Grooms 2016, Allaire et al. 2018). This is an important gap in the literature, especially if the long-run returns to environmental pollution regulation are large, because they have generally been excluded from the benefit-cost analysis of regulation.

In this paper, I investigate how early childhood exposure to lead in drinking water affects both short-run and long-run educational outcomes, providing important information for the economic assessment of the impacts of lead regulation. Additionally, I quantify whether and how lead exposure's impacts differ by socioeconomic characteristics such as gender and race, allowing me to describe the likely distributional impacts of lead regulation and lead mitigation investments.

Using confidential education data from Texas on the educational outcomes for over 2.6 million students and data on drinking water lead contamination from the US EPA via two Freedom of Information Act (FOIA) requests, I use two identification strategies to plausibly causally estimate the educational impact of drinking water lead exposure. In the short run, I employ two instrumental variables (IV) strategies to estimate the impact of lead on third-grade standardized test scores. My first IV strategy exploits the plausibly exogenous variation of surface water chloride (Cl^{-}) concentration that affects the lead concentration in finished drinking water. My second instrument interacts surface water Cl^{-} with the historical presence of lead pipes. I find that even a very low level of drinking water lead exposure in a child's birth year has significant negative impacts on third-grade standardized test scores. Specifically, I find that eliminating lead in Texas drinking water could increase average reading scores by 0.58 percent and math scores by 4.7 percent of a standard deviation. My estimates are in line with standard estimates from the literature using blood lead levels (BLLs), a more precise physiological measure of lead exposure that is not systematically available at a spatial and inter-temporal scale to allow this kind of analysis.

I then employ a difference-in-difference (DID) model using plausibly exogenous variation from the timing of drinking water lead treatment technique violations. Linking the timing of a violation of the federal regulation of drinking water lead treatment with the year of birth, I find that students who experience a new violation in their birth year have a statistically significantly lower probability of graduating high school. Given the income premium of having a high school diploma, a lead treatment violation at birth may be associated with a 0.14 percent decrease in annual income through its impact on high school graduation. For the cohorts in my sample, the benefits in terms of increased income from eliminating such violations in Texas would be around $12 million annually. Additionally, this paper finds that girls, children from African American families and children with economic disadvantages are more vulnerable to lead exposure, suggesting that early childhood lead exposure may be one contributing factor of the gender-achievement gap and racial-achievement gaps in the US.

This paper makes the following contributions to the literature. First, I provide the first evidence of the impacts of contemporary U.S. drinking water lead level on elementary school test scores, showing that exposure even at the low levels typical of regulated U.S. water systems may cause damages. Lead in drinking water is an ongoing public health crisis in some cities and has larger impacts on very young children and pregnant mothers. While the United States (US). has one of the best water supply systems (Columbia Water Center

2016), elevated lead concentrations in drinking water in Flint, MI, Newark, NJ, and Washington DC are among recent high-profile occurrences. Most papers that estimate damages from lead exposure focus on airborne lead from gasoline, and few studies examine the impact of lead exposure from drinking water.

Second, this paper provides the first evidence that the negative effects of early childhood lead exposure persist through longer educational milestones, such as high school graduation. Previous studies have documented lead's impact on children's early cognitive ability, intelligence score (Ferrie et al. 2012), educational outcomes (Reyes 2015*b*, Aizer & Currie 2019), and later crime rate, risky behavior (Reyes 2015*b*) and juvenile delinquency (Aizer & Currie 2019). However, few studies estimate the long-run impacts of lead exposure, such as the impact on higher educational attainment or future earnings.

Third, this paper also contributes to the literature on lead exposure's implications for inequality. Economic and racial inequality can cause poor and minority children to have greater exposure to lead, and prior work suggests that lead may be one cause of continuing disparities in test scores (Aizer & Currie 2019). Studies have shown wide racial and class inequality in lead exposure. From 1999-2010, 7.7% of non-Hispanic black children aged 1–2 years had blood lead levels (BLLs) ≥ 5 $\mu g/dL$, compared with 1.6% of Mexican-American children aged 1–2 years (Raymond & Brown 2017). While existing research finds that differences in housing conditions and exposures to lead-contaminated house dust contribute strongly to the racial disparity in urban children's BLLs (Lanphear et al. 2000), other work reports that the racial disparity in BLLs persists even after controlling for detailed controls such as neighborhood-level education, poverty, and age of neighborhood housing (Sampson & Winter 2016). Moreover, not only are children from minority and poor neighborhoods disproportionately exposed to lead, but they also have less access to mitigating measures such as good nutrition to offset damages caused by lead exposure (Gallicchio et al. 2002, Committee on Environmental Health and others 2005). Following the resurgence of the Environmental Justice (EJ) literature, understanding the causes of inequality in lead exposure and consequences of lead exposure for social inequality could contribute to the discussion on new approaches and policies for reducing inequality (Banzhaf et al. 2019).

The rest of the paper proceeds as follows. In Section 3.2, I review the literature on the health and educational impacts of lead exposure. Section 3.3 provides a brief introduction of the law governing lead concentrations in drinking water in the United States. Section 3.5 describes the data used in this paper, and Section 3.4 presents econometric models. Main results and robustness checks are presented in Section 3.6. In Section 3.7, I include a rough back-of-the-envelope estimation from the benefits of eliminating lead in drinking water, along with conclusions.

3.2 Literature review

Despite more than 30 years of efforts limiting lead exposure, 4.5 million households in the United States are still exposed to high levels of lead, and half a million preschool-aged children have elevated BLLs. The share of U.S. children with elevated BLLs increased slightly from 2009-2014 (Raymond & Brown 2017). Lead contamination disproportionately affects children in minority and low-income families; average BLLs among African American children ages 1-5 in the United States are 38 percent higher, and levels among low-income children are 23 percent higher than the national average (Wheeler & Brown 2013). Using data from more than 1 million blood tests in Chicago, Sampson & Winter (2016) show that disproportionate exposure of lead by race is a pathway through which racial segregation has contributed to racial inequality. Racial and income differences in exposure can be exacerbated by the fact that good nutrition and cognitive stimulation can protect against some negative effects of lead exposure, but children in low-income and minority households may be less likely to benefit from these protective measures (Benfer 2017).

The three main historical sources of lead in the U.S. are lead paint, leaded gasoline, and lead in drinking water (Reyes 2015*a*). Many laws and regulations have been passed to restrict lead use in commercial products. The phase-down of lead in gasoline began in 1974 under the authority of the Clean Air Act of 1970 (Newell & Rogers 2003). By 1996, lead was banned as a fuel additive in the United States. Thus, the only remaining sources of airborne lead from gasoline in the United States are the vehicles in which leaded fuel is still legal, professional auto racing (Hollingsworth & Rudik 2020) and aircraft (Zahran et al. 2017). Federal regulations banned the use of lead-based paint in homes built after 1978, but lead paint remains the most common source of lead exposure in childhood.

The U.S. Environmental Protection Agency (2017) estimates that lead in drinking water contributes to 20% or more of a U.S. individual's total exposure to lead. Lead in drinking water has a larger impact on very young children and pregnant mothers. For example, infants who consume mostly mixed formula can receive 40%-60% of contemporaneous lead exposure from drinking water (U.S. Environmental Protection Agency 2017). EPA promulgated the Lead and Copper Rule (LCR) under the Safe Drinking Water Act in 1991. The LCR requires public water systems to monitor lead concentrations at customer taps. If observed lead concentrations exceed an action level of 15 parts per billion (ppb), water utilities are required to undertake several actions to control corrosion, which reduces the leaching of lead from water pipes and fixtures (U.S. EPA 2000).

Early childhood exposure to lead has irreversible health and behavioral consequences (Hanna-Attisha et al. 2016), and it is particularly dangerous for fetuses and young children, who absorb lead more efficiently than adults and are also at stages of rapid development of the neurobehavioral system (Bellinger 2008). The public health and epidemiology literature suggest that lead affects children's brain development, resulting in reduced intelligence quotient (IQ), increased likelihood of attention deficit hyperactivity disorder (ADHD), behavioral changes such as reduced attention span and increased antisocial behavior, and unsatisfactory educational outcomes (Schwartz 1994, Chen et al. 2007, Bellinger 2008, Aizer et al. 2018).

Economists and sociologists also have documented near-term social and behavioral impacts of lead exposure outside of the educational system. Ferrie et al. (2012) estimate the negative impacts of water-borne lead on World War II Army enlistees' cognitive ability and IQ. Others have documented the impact of lead on criminal and other risky behavior in and after adolescence (Reyes 2015*b*), as well as juvenile delinquency (Aizer & Currie 2019, Sampson & Winter 2018).Other studies focus on infant mortality (Clay et al. 2014*a*), IQ (Ferrie et al. 2012) and violent behavior (Feigenbaum & Muller 2016).

But extrapolating the effects of short-term impacts (IQ and education outcome in third grade) to long-term impacts (higher education participation, employment, and earnings) is challenging for two reasons. First, the relationship between primary school test scores or IQ and educational attainment is unlikely to be linear. So marginal effects from studies linking lead exposure to these intermediate outcomes cannot easily predict impacts on educational attainment and eventual labor-market outcomes. Second, many existing studies that estimate lead's intermediate impacts use BLL as the key independent variable. But good nutrition can reduce lead absorption, making BLL a noisy measure of absorption. In addition, absorption is itself, a noisy measure of damages. Lead damages may accrue over years of childhood exposure, mitigated to differing degrees by nutrition, cognitive stimulation, and other parental and school inputs, so estimates of marginal effects of pre-school BLL on early grade test scores (for example) may not provide an accurate picture of eventual educational attainment in later years. Moreover, the differential impact of lead exposure may affect educational attainment across generations.

Thus, studies that quantify the impacts of lead exposure from drinking water at current (non-historical) levels on long-run outcomes are desirable. However, to my knowledge, the only prior work that is Grönqvist et al. (2020). They use linked individual data in Sweden and find that lead phaseout from gasoline is associated with 4% increase in future income annually.

3.3 The Lead and Copper Rule

The regulation governing lead concentrations in U.S. drinking water is the Lead and Copper Rule (LCR). The LCR was promulgated by EPA in response to the 1986 SDWA Amendments and regulates lead contamination at households' taps. It sets a Maximum Contamination Level Goal (MCLG) for lead of 0, which means there is no safe level of lead in drinking water, and any amount is considered harmful to human health. Unlike many other well-studied regulations under the SDWA, the LCR is a treatment technique rule,, without an enforceable maximum contamination level (MCL). It specifies a list of treatment, monitoring, and public education guidelines to ensure that water systems provide safe water to their customers, requiring water systems to sample water from the taps with a higher chance of having lead in drinking water twice every 6 months and measure the lead concentration. A series of actions are triggered when the fraction of samples exceeding 0.015 mg/L of lead is found to be greater than 10 percent. Actions include examining source water quality, installing a state-designated corrosion control treatment (CCT) program, public education, and lead service line removal. All the required actions are aimed to reduce the amount of lead leaching into the water from different plumbing materials. If a public water system violates the required monitoring, treatment, and public education guidelines, it triggers an LCR violation.

The LCR is one of the most complicated drinking water regulations for states and drinking water utilities to implement due to the need to control corrosivity of treated drinking water as it travels through distribution systems to the consumer's tap (US Environmental Protection Agency 2016). States and public water systems must have expertise and resources to identify the sampling locations and to collect and analyze samples correctly. They also need more resources to identify and install the optimal CCTs and maintain the effective operation of the CCTs. Given the requirements in resources and expertise to comply with LCR, small systems that may lack these resources are more likely to have LCR violations. Using data from U.S. EPA's Safe Drinking Water Information System (SDWIS), the utilities that experience LCR violations serve about 3500 people on average. 90% of systems that experienced LCR violations serve fewer than 10,000 people, which is the EPA upper threshold for small systems(USEPA 2012).

The LCR has two main types of violations: monitoring and reporting violations, and treatment technique violations. Monitoring and reporting violations occur when public water systems fail to monitor and report water quality in a timely manner. It includes when they fail to test both lead concentration at consumers' taps and test water quality parameters, such as pH and alkalinity, in the source water. Treatment technique violations include failing to submit an optimal corrosion control technique (OCCT) study or recommendations, failing to install OCCT on time, not meeting water quality parameter requirements, having lead or copper above a state-designated permissible level, not replacing lead service lines, or failing to send out public education materials. It is important to note that having lead concentration over the action level does not trigger a violation. Violations only occur when public systems fail to follow designated treatment techniques and their associated timelines. Under the SDWA, PWSs can violate drinking water standards by having health-based violations, monitoring and reporting violations, or public notice violations (US EPA 2019). The health-based LCR violations are more serious since they mean systems fail to follow the treatment technique requirements. However, the non-health-based violations, such as monitoring and reporting violations can also pose threats to drinking water quality (Fedinick et al. 2017). We may not know about the exceedance in lead concentration when a PWS fails to monitor water quality.

Since the protocols are set to ensure that public water systems (PWSs) minimize lead in drinking water, there are good reasons to believe that violation of these protocols is associated with a potential threat to public health. Even though violations may not indicate the presence of lead in a system's drinking water, they may be correlated with increased lead concentrations. For instance, when a public water system does not report a lead concentration to the appropriate state agency, high lead levels can go unobserved. During the Flint water crisis, Flint was not listed as having an LCR violation in the EPA violation database in 2015 but its underlying lead concentration was already high (Olson & Fedinick 2016). According to a study of 72 Flint households in August 2015, 20% of the samples had lead levels that exceeded the action level (0.015mg/L), and the 90th percentile was 30 $\mu g/L$ (Masten et al. 2016).

Table 3.1: The association between lead concentration and LCR violation

	(1) LCR	(2) LCR	(3) LCR
Lead concentration	17.93*	30.18*	32.75**
	(9.258)	(17.08)	(14.16)
Year FE	No	Yes	Yes
County FE	No	No	Yes
Observations	6900	2497	1949
R^2	0.0018	0.0113	0.0610

Robust standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To demonstrate this correlation, consider the coefficients reported in Table 3.1, which shows results from a logistic regression of LCR violations in 3010 Texas public water systems from 2006-2011 on observed lead concentrations as measured in mg/L, controlling for year and county fixed effects. The results indicate that lead concentration is positively associated with a PWS having an LCR violation[1]. Also it is important to note that once in violation, it takes the PWSs years to be back in compliance. As is shown in Table 3.2,

[1]Because the LCR status of some counties does not change during the time, adding county and year fixed effects shrinks the sample size.

the average duration of a violation is 1115 days, more than 3 years.

The LCR violation and current LCR rule could also pose environmental justice issues. Lead service lines (LSLs) are the most important source of lead from drinking water. But under the LCR, water systems are only accountable for the public portion of the LSL replacement, leaving on average $3,000 dollars to homeowners(US EPA 2021*b*). Past studies have found strong correlation between full LSL replacement and family income and race using data from Washington DC (Environmental Defense Fund 2020). Moreover, when corrosion control alone is not sufficient, consumers need to take further actions to reduce their exposure to lead, such as installing water filters or switching to bottled water. Consumers' ability to understand and afford these actions pose additional challenges to low income families.

3.4 Empirical Strategy

I identify the effect of early childhood exposure to lead from drinking water on school-aged and young adult outcomes by exploiting variation in the magnitude and timing of exposure among individuals born between 1992-2001 and 2006-2011. The amount of lead exposure experienced by an individual depends on their year and county of birth. The identifying assumption is that after flexibly controlling for many observable and unobservable potential confounders, changes in a county's drinking water lead concentration at the birth year affects the educational outcome of individuals born in the particular county in later years.

3.4.1 Baseline Econometric Model

Following Isen et al. (2017), my baseline econometric model is an OLS regression:

$$Outcome^{a}_{ict} = \beta_0 + \beta_1 logLead_{ct} + \mathbf{X}_i\theta + \mathbf{N}_{ct}\psi + \gamma_c + \alpha_t + \varepsilon_{ict} \qquad (3.1)$$

Because I only have lead concentration data from 2006-2011, I use the baseline OLS regression to examine the impact of lead on individuals born during this period on their short-term outcomes. $Outcome^{a}_{ict}$ is the scaled 3rd grade reading or math score, or a categorical variable indicating whether the child failed, met or mastered one of these tests. $logLead_{ct}$ is the log transformation of average lead concentration from drinking water of an individual's birth county in the year of birth. Because the response function for lead is a "hockeystick shape" and unlikely to be linear, I use log-transformed lead in my analysis (Grönqvist et al. 2020). $\mathbf{X}_i$ is a vector of individual characteristics including gender, race and economic disadvantage status. $\mathbf{N}_{ct}$ is a vector of county-level, time-varying characteristics, including median income, percentage of people below the poverty line and the unemployment rate. γ_c is a birth-county fixed effect that controls for time-invariant, unobserved characteristics that could affect 3rd grade standardized test scores for individuals born in a particular county. α_t is a birth-year fixed effect that controls for time-varying determinants of standardized test scores that are common to all individuals born in Texas a given year. By using these fixed effects, we are comparing individuals born within the same county in different years. $\hat{\beta}_1$ is the coefficient of interest and measures the effect on third-grade test scores of an increase in the lead concentration in drinking water in the birth year.

The OLS model assumes that the unobserved determinants of test scores should not covary with changes in lead exposure conditional on the covariates. However, this assumption can be violated if there exist any unobserved determinants that also change with lead exposure over time by county. For instance, neighborhoods with higher level of lead may be in older cities and have higher population density. If unobservable factors change differently overtime between old and newer cities, or between urban and rural areas, the OLS model coefficient estimates are likely to be biased.

Moreover, people from certain racial groups and economically disadvantaged families are more likely to be exposed to higher levels of lead from drinking water (Banzhaf et al. 2019, Marcus 2020). Though I use birth-county fixed effects to capture time-invariant characteristics, this identification strategy may still suffer from omitted variable bias. Locations with different levels of lead may also have different underlying conditions such as average income or crime rates. People with different backgrounds or preferences for clean water might sort into locations with varying levels of lead concentration (Deschenes & Meng 2018). To address concerns about endogenous lead exposure, I use the three identification strategies including the plausibly exogenous variation from Cl^- levels in surface water and the presence of lead pipes in 1900 as instruments (Clay et al. 2014*a*, Stets et al. 2018), and the exogenous timing of LCR violations.

To analyze the endogenous exposure to lead, I use the directed acylic graph (DAG) (Figure 3.1) to explain my identification strategy. DAG is a graphic presentation of the causal effects using nodes and arrows (Cunningham 2021). Nodes represent random variables created by some data-generating process and arrows represent a causal effect between two random variables moving in the direction of the arrow. In Figure 3.1, I have a list of variables including: drinking water lead exposure at birth, educational outcomes, lead pipes, water treatment, source water chemistry, geology, neighborhood characteristics (public funding, neighborhood resources, urbanization), and individual characteristics (family income, race and ethnicity). These variables are connected by arrows representing the causal relationship among them.

There are a list of direct and indirect paths between drinking water lead exposure and later educational outcomes:

1. Drinking water lead at birth $\longrightarrow$ Educational outcome (direct path 1)
2. Drinking water lead at birth $\longleftarrow$ Lead pipes $\longleftarrow$ Family income, race and ethnicity $\longrightarrow$ Educational outcomes (backdoor path 1)
3. Drinking water lead at birth $\longleftarrow$ Lead pipes $\longleftarrow$ Neighborhood characteristics (public funding, resources and urbanization) $\longleftarrow$ Family income, race and ethnicity $\longrightarrow$ Educational outcomes (backdoor path 2)
4. Drinking water lead at birth $\longleftarrow$ Water treatment $\longleftarrow$ Neighborhood characteristics (public funding, resources and urbanization) $\longleftarrow$ Family income, race and ethnicity $\longrightarrow$ Educational outcomes (backdoor path 3)

5. Drinking water lead at birth $\longleftarrow$ Water chemistry $\longleftarrow$ Urbanization $\longleftarrow$ Family income, race and ethnicity $\longrightarrow$ Educational outcomes (backdoor path 4)

6. Source water chemistry and geology $\longrightarrow$ Water chemistry $\longrightarrow$ Drinking water lead at birth $\longrightarrow$ Educational outcomes (direct path 2)

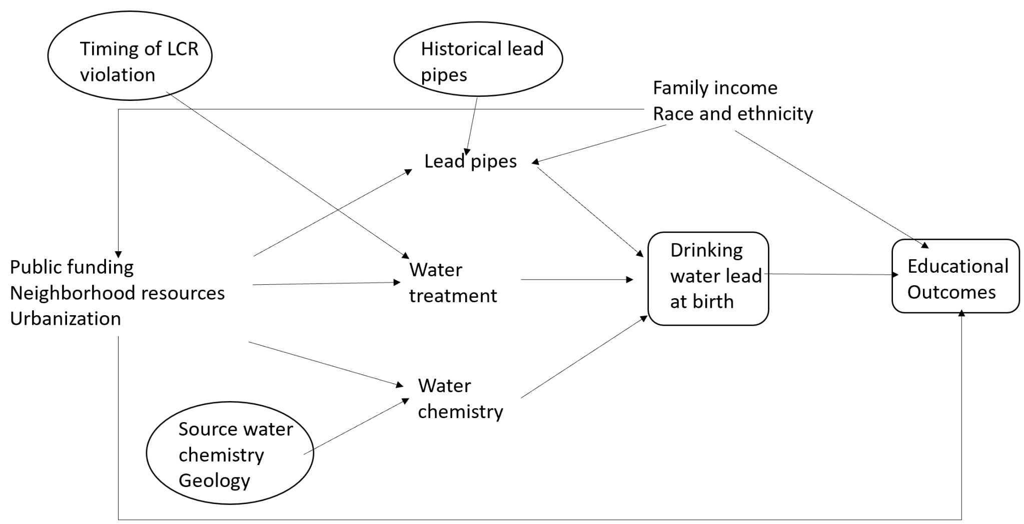


Figure 3.1: DAG of drinking lead impacts

Notes: Figure shows the DAG of drinking water lead impact on educational outcomes.

There are 6 paths between drinking water lead exposure at birth and later educational outcomes. The first path is the direct causal relationship between drinking water lead exposure at birth and educational outcomes. Paths 2 through 5 are backdoor paths, which means that the correlations between lead exposure and educational outcomes are driven solely by fluctuations in variables such as family income and neighborhood resources. Path 6 is another direct path from drinking water lead exposure to educational outcomes. This is different from the first direct path because this path is driven by changes in source water chemistry and geology, which are unrelated to backdoor variables.

Lead enters drinking water when water corrodes pipes and fixtures that contain lead (US EPA 2021*a*). There are three crucial parts to the process. First, there need to be plumbing materials containing lead, such as pipes, solders, or fixtures. Second, water utilities are required to implement optimal corrosion control techniques, such as adjusting pH or adding orthophosphate to make lead-phosphate scale in pipes. But these approaches may not always be useful and could have unintended consequences given the complexity of water chemistry in pipes and new findings from environmental engineering research (Pelley 2018). Moreover, water utilities need to test for drinking water lead concentration. Failing to meet the monitoring schedule at the water utility level could cause lead pollution to go undiscovered. Third, water corrodes lead pipes when drinking water is acidic or contains disinfectant, inorganic carbon, iron, manganese, and aluminum compounds, or other components that promote the corrosion of scale in lead pipes and cause a release of lead particles. While water utilities alter the water chemistry, factors such as temperature and weather may influence the chemistry of drinking water even after utilities implement corrosion control techniques. (Roy & Edwards 2019).

The backdoor paths in Figure 3.1 show the complex relationship between various socioeconomic variables and the factors affecting drinking water lead exposure. Backdoor path 1 shows that the family background of an individual, such as parents' income, race, and ethnicity, affects the probability that an individual is born in a house with lead pipes. As mentioned before, replacing lead pipes could be expensive to many low-income residents even with federal and state funding programs. For example, the Trenton (New Jersey) Water Works Lead Service Line Replacement Program limits homeowner expenses to $1,000 by covering the rest of the cost for replacement, which typically runs between $3,000 and $7,000 (Santucci & Scully 2020). Family income could also affect students' educational outcomes by investment in students' education and productivity.

Backdoor path 2 suggests that on top of backdoor path 1 relationships, family income, race, and ethnicity could also sort people to live in certain neighborhoods. While the LCR has a requirement on the lead service line replacement rate, it is up to communities and utilities to make plans for this replacement. Wealthy neighborhoods may have more resources to replace lead pipes. There are also better schools in those neighborhoods that would lead to better student educational outcomes. Backdoor path 3 and backdoor path 4 show a similar story. Water districts and communities with more funding and resources could implement better corrosion control techniques. Urbanization, though, can also influence drinking water chemistry and lead to a higher level of lead from drinking water.

Since there are 4 *open* backdoor paths, I control for family income, race and ethnicity, neighborhood income, population, and poverty to remove bias. Also, I use exogenous variation in surface water chemistry, the presence of historical lead pipes, and the timing of LCR violations to discern a causal relationship between drinking water lead exposure and educational outcomes.

3.4.2 Instrumental variables models

Prior research has shown that surface water trends in Cl^- affect corrosivity in water distribution systems, which further affects drinking water quality (Stets et al. 2018). I use two different instruments that exploit this fact: the level of Cl^- in source water in an individual's birth county-year, and the interaction of surface water Cl^- and the presence of lead pipes in an individual's birth county in 1900. Surface water Cl^- level is a likely valid instrument for the following reasons. First, Cl^- is strongly related to elevated lead concentration in the drinking water. Past studies have found that high Cl^- concentration promotes galvanic corrosion of materials containing lead, such as lead service lines and lead solder in water distribution systems (American Water Works Association 1996, Edwards & Triantafyllidou 2007, Willison & Boyer 2012, DeSantis et al. 2018), a water chemistry issue that played a role in the ongoing crisis in Flint. [2] Similarly, in Washington DC from 2004-2006, where the city's PWS switched from free chlorine to chloramine as a disinfectant, the failure to control corrosion caused elevated lead levels(Edwards et al. 2009).

[2] When the water authority in Flint switched its water source from Lake Huron to water from the Flint River, the high level of Cl^- in Flint river was not properly treated to optimize corrosion control, and Flint water ended up having observed lead concentrations over 5000 ppb, way above the action level of 15 ppb (Pieper et al. 2017, 2018, Torrice 2020).

Second, surface water Cl^- concentration is strongly correlated to Cl^- level in finished drinking water. Cl^- is a naturally occurring ion and mostly exists in the form of sodium chloride in water. There are no federal or state primary health-based drinking water standards for Cl^-, but only an advisory standard for the aesthetic purpose (e.g. taste) (U.S. Environmental Protection Agency 2003). Though it is possible that people drink less water if the water is too salty from a high level of Cl^-, the average level of Cl^- in my data is less than the standard that the US EPA identifies as a concentration that Cl^- can be expected to cause a salty taste in drinking water (New Hampshire Department of Environmental Services 2010). Moreover, typical drinking water treatment techniques, such as alkalinity or pH adjustments, do not remove Cl^- and drinking water treatment plants do not desalinate, and therefore high Cl^- concentrations should be conserved even after treatment (Stets et al. 2018). It is important to note that I'm using the surface water Cl^- concentration, because finished drinking water Cl^- source are not available.

Third, Cl^- is not likely to be correlated with educational outcomes other than by affecting lead concentrations. Cl^- is an essential nutrient for human health. But drinking water would typically contribute only 2.5% to 5% of the dietary salt goal if tap water consumption is 2 L/day (U.S. Environmental Protection Agency 2003). Since the intake of Cl^- from drinking water is only a small portion compared to other pathways, it is not likely to affect human health or educational outcomes.

The identifying assumption is that after controlling for individual and neighborhood characteristics and fixed effects, changes in surface water Cl^- of a county are unrelated to the changes in education outcomes of people born in the county later in life except through the changes in drinking water lead concentration. One threat to my identification is that the change in the Cl^- across counties over the years could relate to factors that would affect education outcomes that my models do not control for. For example, anthropogenic sources of Cl^- include but are not limited to the use of fertilizers and irrigation in agricultural fields, road salts, and wastewater discharge or runoff from urban areas (Panno et al. 2006, Steele & Aitkenhead-Peterson 2011). As a result, the change in surface water Cl^- could be correlated with the change in the degree of urbanization, which may be correlated with educational outcomes later in life. However, I believe this is less of a concern given that Cl^- concentrations in source water are also increased by the use of irrigation and fertilizers in agricultural fields. Using data on surface water quality from 1982–2012, Stets and coauthors find that freshwaters are being salinized rapidly in all kinds of landscapes in the U.S. (Stets et al. 2020). Moreover, one major source of Cl^- in the U.S. is from the use of road salts for melting snow, which is less common in Texas.

I also use the interaction of birth-year surface Cl^- level and the presence of lead pipes in 1900 as an additional instrumental variable. For each individual, I use the data from Clay et al. (2014*b*) on lead pipes to determine if her water system historically had lead pipes. While Clay et al. (2014*b*) data only has information on the presence of lead pipes in 172 large and medium-sized cities, I geocode the cities to the counties of location. My instrument is not the presence of lead pipes but the interaction of surface water Cl^{-} and the historical presence of lead pipes. While the temporal and spatial variations in Cl^{-} level may be affected by the changes in the degree of urbanization and agricultural activity over time and across counties, having lead pipes in 1900 is correlated with the contemporary presence of lead pipes but less likely to be correlated with the urbanization and agricultural activities of counties during my sample period. The identifying assumption is that Cl^{-} level in surface water is more predictive of lead concentration changes where there is a historical presence of lead pipes.

The first-stage regression in the two-stage least square (2SLS) estimator is as follows:

$$Lead_{ict} = \alpha_0 + \alpha_1 Z_{ict} + \mathbf{X}_i\theta + \mathbf{X}_{ct}\psi + \gamma_c + \eta_t + \varepsilon_{ict} \tag{3.2}$$

where $Lead_{ict}$ indicates the lead concentration in drinking water. In the first IV, I use an indicator of lead concentration action level exceedance in a county c and year t when an individual i was born. In the first IV, I regress lead on the weighted average Cl^{-} concentration of surface water in an individual's birth county and birth year. I first calculate the average Cl^{-} level of a watershed at a given year. Then I estimate the average Cl^{-} concentration of a county using the overlapping area between a county and a watershed as weights. In the second IV, I use the concentration of drinking water lead as the dependent variable and use the interaction of surface water Cl^- and lead pipes as the Z_{ict}.

In the second stage, I use the predicted indicator of lead from equation (3.2) in the place of actual lead concentration.

$$Outcome_{ict} = \delta_0 + \delta_1 \widehat{Lead_{ict}} + \mathbf{X}_i\theta + \mathbf{N}_{ct}\psi + \gamma_c + \eta_t + \varepsilon_{ict} \qquad (3.3)$$

The coefficient of interest in equation (3.3) is δ_1, which measures the effect of having lead concentration over the SDWA action level in an individual's birth county-year on her 3rd-grade standardized test scores in the first IV. The second IV measures the effect of lead concentration in an individual's birth county-year on her 3rd-grade standardized test scores.

3.4.3 Using LCR violations for long-run outcomes

While the IV approach can solve potential endogeneity problems, it is constrained by the availability of lead concentration data, so I cannot use it to study the effect of lead on long-run outcomes, such as the high school graduation rate. Thus, for longer-run outcomes, I use plausibly exogenous variation from the timing of water quality violations (Currie et al. 2013, Marcus 2020) and employ a fixed-effects model to estimate impacts of childhood lead exposure from drinking water on individuals in Texas. I explore the effect of LCR violations on individual outcomes using the following specification:

$$Outcome_{ct} = \beta_0 + \beta_1 LCR_{ict} + \mathbf{X}_i\theta + \mathbf{N}_{ct}\psi + \gamma_c + \eta_t + \varepsilon_{ict} \tag{3.4}$$

where the outcome is the average standardized test score in 3rd grade, or the likelihood of passing the standardized tests for the younger cohorts and high school graduation rate for the older cohorts. LCR_{ct} measures if the cohort experiences a new LCR violation in their birth county c in the birth year t. Similar to Equation 3.1, I also control for individual characteristics, county-of-birth fixed effects and year-of-birth fixed effects. With the fixed effects, I compare children born in counties without a new LCR violation to children born exposed to a new LCR violation. β_1 is the coefficient of interest, which measures the effects of exposure to LCR violations on the outcome of interest. Standard errors are clustered by county.

One key advantage of this approach is that it allows me to use data for all the cohorts from 1992-2011 and examine the impact of lead on long-run outcomes, such as high school graduation. The assumption for this specification is that the variation in LCR violation exposure is "as good as random". The assumption is likely to hold with the county fixed effects for the cross-sectional difference between counties that might be related to LCR violations and the educational outcomes of children. For instance, individuals born in poor counties with worse public schools may also have a higher probability of exposure to LCR violations since the PWS may have less funding to test lead concentrations in a timely manner. However, in the presence of time-varying unobservable confounders, my estimate will be biased. For example, LCR violations may happen because a PWS's funding is reduced during the financial crisis, which could also affect the children's health and educational outcomes. If this is the case, I may overestimate the impacts of LCR violations.

Thus, I use a difference-in-difference estimator for this specification. As mentioned before, the LCR is a treatment technique rule. All the technical requirements in the LCR are designed to reduce the likelihood of lead exposure in drinking water. From the IV models, I estimate the impacts of lead from drinking water. The difference-in-difference estimator with this specification evaluates the effectiveness of the policy that aims to reduce lead levels from drinking water.

3.5 Data

I use data from multiple sources. Educational outcome data come from the Texas Education Research Center (ERC), which provides linked individual-level education and workforce administrative data. Drinking water quality data comes from the EPA's Safe Drinking Water Information System (SDWIS) data, and EPA's National Contaminants Occurrence Database (NCOD). Surface water quality data is obtained from the Water Quality Portal from the U.S. Geological Survey (USGS) and the Surface Water Quality Monitoring program of the Texas Commission on Environment Quality (TCEQ). The U.S. Centers for Disease Control and Prevention (CDC) provides measures of children's BLLs. I also use data from other administrative surveys, such as the American Housing Survey, to obtain estimates for population, household income, unemployment rates, poverty levels, and other community characteristics.

3.5.1 Texas education and income data

I use administrative data from the Texas Education Research Center (ERC) for my outcome variables. The Texas ERC collects student-level data from the Texas Education Agency (TEA) and the Texas Higher Education Coordinating Board (THECB). It also links students to their workforce participation data from the Texas Workforce Commission (TWC). The data include all students who ever enrolled in public schools in Texas and cover enrollment, standardized test scores, disciplinary action records, high school graduation status, public secondary school enrollment, quarterly earnings, and basic demographics such as race, gender, and economic disadvantage status.

To match the inter-temporal availability of lead concentration data, I use individuals born in Texas from 2006 to 2011 for whom I have 3rd-grade standardized test scores to estimate short-term impacts, and individuals born in Texas from 1991 to 2001 who ever enrolled in 9th grade in Texas to estimate long-term impacts. A key limitation is that the data do not identify an individual's date of birth or place of birth. The TEA enrollment data has the year of enrollment, grade of enrollment, and age as of September 1st. I identify the list of individuals who enrolled in early education (EE), prekindergarten (PK), or kindergarten (KG) in Texas and estimate their year of birth using year of enrollment and age as of September 1st. I refer to students using their estimated year of birth, so my first cohort is students who were born in 1989. The TEA data has the county in which a school is located, and I use as a proxy that the county where an individual enrolled in EE, PK, or KG is her county of birth. I also obtained the detailed location (latitude and longitude) of public schools from the Common Core of Data from the National Center for Education Statistics. Of the 11,849 public K-12 schools in Texas, 181 of those schools do not have location information. Thus, I assume the latitude and longitude of enrolled schools as the location of birth as a robustness check.

I consider several outcome variables. First, I use scaled reading and math standardized test scores in third grade. My study period includes a major change in Texas standardized testing, the 2012 shift from the Texas Assessment of Knowledge and Skills (TAKS) to the State of Texas Assessments of Academic Readiness (STAAR). For cohorts born between 2006 to 2011, I use their STAAR scores. For cohorts born between 1989 to 1994, I use their TAKS scores[3].Because the STAAR and TAKS are different tests, I do not merge them together and compare outcomes. I also use flags indicating whether an individual passes the relevant standardized tests as a binary outcome variable.

Table 1 reports summary statistics, with the the IV sample in Panel A and the DID sample in Panel B. The IV models use 3rd-grade test scores from more than 1 million students born between 2006-2011. The average

[3]STAAR is designed to measure students' ability to apply the knowledge and skills defined in the state-mandated curriculum standards, the Texas Essential Knowledge and Skills. STAAR was first administered in 2012 and is the current standardized test used in Texas. TAKS is the test implemented in Texas before STAAR, from 2003 to 2011.

scaled reading score is 1431 points with a standard deviation of 173 points. The average scaled math score is 1465 points with a standard deviation of 174 points. I use the scaled test scores because the scaled scores consider the difficulty level of each individual test question and allow me to compare test results from year to year (Texas Education Agency 2015). About 40% of the students born between 2006-2011 passed the reading standard in 3rd grade and 44% students passed the math standard in 3rd grade. The average age of students enrolled in 3rd grade in Texas is 8 years old.

Second, I examine two different long-run outcomes - high school graduation and enrollment in Texas public universities for the early cohorts. Panel B of Table 3.2 reports the summary statistics for the sample used in the DID model. There are about 1.3 million students born in Texas enrolled in Texas high school from 1991-2004. 78.4% of these students graduate high school but only 25.8% of these students enrolled in Texas public universities. Of this sample, 48.9% are female students. 47.6% of these students are Hispanic and 35.4% of these students are white. 38.7% students receive free lunch, 8% students receive reduced lunch and, 7% have other economic disadvantages.

3.5.2 Water system and drinking water quality data

Data on public water systems and water quality violations comes from the Safe Drinking Water Information System (SDWIS) from the U.S. EPA, obtained via FOIA request by prior researchers (Baker et al. 2019) and shared

Table 3.2: Summary statistics

	count	mean	median	sd	min	max
IV sample						
Reading score	1318783	1431.163	1427.000	174.506	100.000	4011.000
Math score	1318783	1464.782	1458.000	175.466	111.000	4378.000
Meet reading standard	1318783	0.404	0.000	0.491	0.000	1.000
Meet math standard	1318783	0.446	0.000	0.497	0.000	1.000
ALE	1318783	0.022	0.000	0.148	0.000	1.000
Lead concentration (ppb)	1318783	1.856	1.700	1.138	0.000	12.729
Cl concentration (mg/L)	1318783	190.691	73.910	473.370	7.973	23032.160
CSMR	1318783	1.704	1.179	1.309	0.059	23.393
Presence of lead pipes	1318783	0.021	0.000	0.144	0.000	1.000
Female	1318783	0.496	0.000	0.500	0.000	1.000
Native American	1318783	0.006	0	0.075	0	1
Asian	1318783	0.037	0	0.188	0	1
Black	1318783	0.129	0	0.335	0	1
Hispanic	1318783	0.558	0	0.5	0	1
White	1318783	0.253	0	0.435	0	1
Free lunch	1318783	0.466	0	0.499	0	1
Reduced lunch	1318783	0.068	0	0.251	0	1
Other economic disadvantage	1318783	0.132	0	0.338	0	1
Age	1318783	8.088	8.000	0.278	8.000	9.000
Unemployment rate	1318783	6.541	6.700	2.003	1.900	15.300
Poverty rate	1318783	17.913	17.000	7.074	6.000	39.900
Median household income	1318783	48757.090	47159.000	12353.890	23096.000	83968.000
Population	1318783	1591577	779213	1513040	272	4179796
Maximum temperature	1318783	26.433	26.273	1.615	21.351	30.991
Precipitation	1318783	2.747	2.615	1.081	0.294	5.672
DID sample						
High school graduate	1341729	0.784	1	0.412	0	1
Enroll in public university	1341729	0.258	0	0.438	0	1
LCR violation status	1341729	0.195	0	0.396	0	1
Female	1341729	0.489	0	0.5	0	1
Native American	1341729	0.004	0	0.061	0	1
Asian	1341729	0.022	0	0.145	0	1
Black	1341729	0.145	0	0.352	0	1
Hispanic	1341729	0.476	0	0.5	0	1
White	1341729	0.354	0	0.478	0	1
Free lunch	1341729	0.387	0	0.487	0	1
Reduced lunch	1341729	0.082	0	0.274	0	1
Other economic disadvantage	1341729	0.069	0	0.254	0	1
Birth year	1341729	1993.691	1993	2.016	1991	2001
Unemployment rate	1341729	7.33	6.3	4.234	0.9	39.3
Poverty rate	1341729	19.514	17.75	8.27	3.5	52.55
Population	1341729	1011788	1109330	424312	331	3471291
Median household income	1341729	31426.83	32071.5	8109.405	11269.5	77303
Maximum temperature	1341729	25.593	25.464	1.639	19.75	31.674
Precipitation	1341729	3	2.977	2.553	3.981	19.219

with me. The lead concentration data is from the 3rd Six-Year Review pollution occurrence data in the National Contaminant Occurrence Database (NCOD), also obtained via FOIA request.

3.5.2.1 SDWIS data

SDWIS provides the public water system (PWS) identifier, the PWS name, number of people served, and dates of the beginning and the ending of each violation. I focus on violations in public community water supply systems (CWS) in my study. CWSs are defined as the public water systems that supply water to the same population year-round. Out of 15,736 PWSs in Texas, 49% (7,713) are CWSs. Using the methods in Baker et al. (2019), I match LCR violations with the PWS inventory from EPA for PWS characteristics, such as name and state served. I also merge LCR violations with the Government Performance and Results Act from Fiscal Year 2013 to identify the county each PWS serves. Given that the FOIAed SDWIS data obtained from Baker et al. (2019) has violations through 2014, I use the current SDWIS database available from the EPA website to update the data to the second quarter of 2020 and estimate the duration of violations.

I focus on LCR violations that occur in 2006-2011 for the younger cohorts and violations between 1991-2004 for the older cohorts to measure individuals' early childhood lead exposure. I generate an indicator variable of having a new LCR violation in county c in year t and link the indicator variable with each individual's birth county and birth year to define early childhood exposure. During the periods of interest, 1336 new LCR violations happened in CWSs in Texas, affecting 5,675,603 individuals. Figure 3.3 shows the distribution of violations by year. Most Texas LCR violations happened in 1992, 2010, and 2011. By the second quarter of 2020, 1236 violations had returned to compliance. The average duration of a violation is 1115 days, with a standard deviation of 788 days. The average population served by CWSs with LCR violations is about 3700. There is a large variation among counties in terms of the number of LCR violations. On average, each county experienced 60 violations of the LCR from 1991-2011, with a standard deviation of 91. Each public water system had an average of 1.6 LCR violations during the same period.

I use all the LCR violations in this paper including both health-based and non-health-based violations. Among the 1336 violations in Texas, 1305 violations are monitoring and reporting violations (non-health-based violations) and 31 violations are health-based violations. In my sample of the older cohort, 19.5% of students who enrolled in high school were born in counties with new LCR violations in their birth year.

3.5.2.2 NCOD data

The lead concentration data come from the 3rd Six-year Review of pollution occurrence data from NCOD. The SDWA requires EPA to review each National Primary Drinking Water Regulation (NPDWR) at least once every six years. During the Six-Year Review process, EPA analyzes SDWA

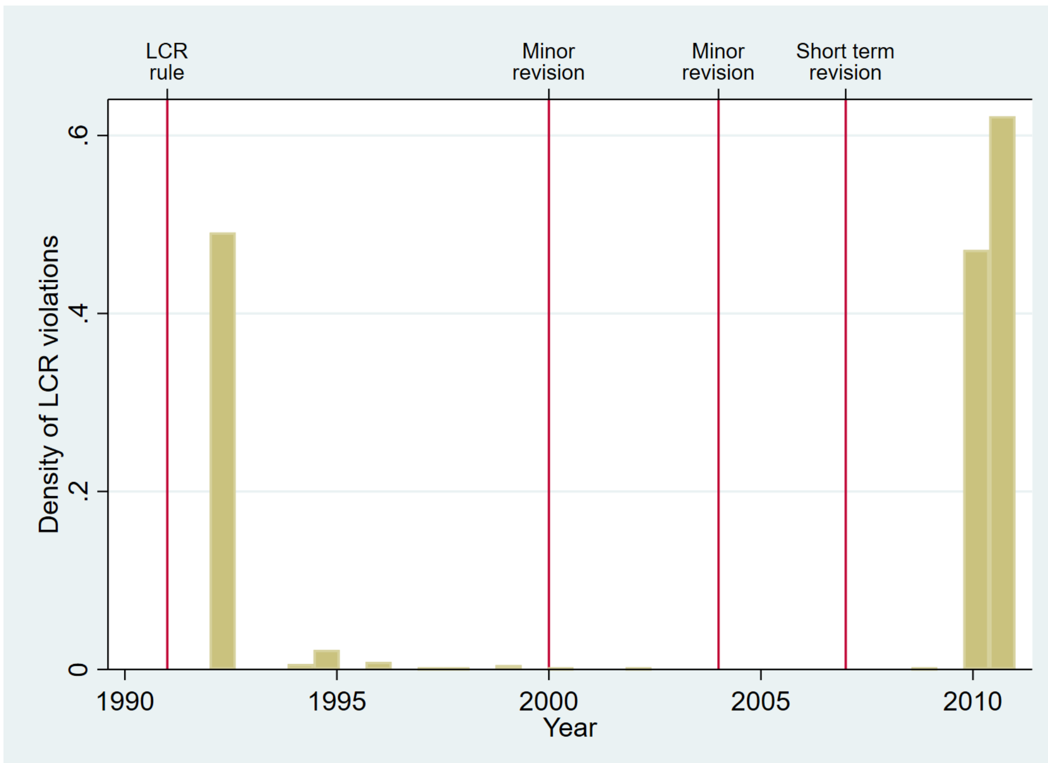
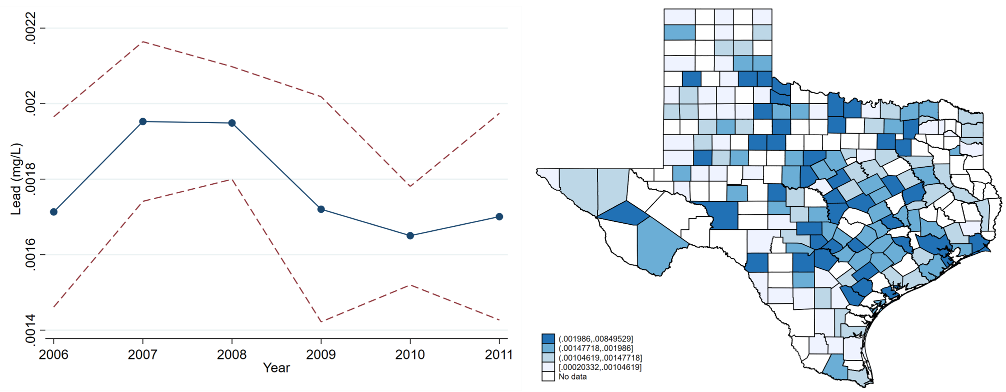


Figure 3.2: Number of LCR violations by year, 1992-2011

Notes: Figure shows the distribution of LCR violations by year. The red lines show years when changes happen to the Lead and Copper Rule.

compliance monitoring data from public water supplies for regulated drinking water contaminants and publishes review results. The NCOD collects data used for the review process. The NCOD data provide information on the public water system ID, lead concentration, detection limit, and the date of testing. I use the date of testing to merge with a cohort's birth year. Given that the NCOD data do not have information on the service area of PWSs, I use the geographic area data from the SDWIS to define the county each PWS serves. The NCOD reports two lead measures. One lead measure reports 0 if the lead concentration is lower than the detection limit of the testing technique, while the other measure reports a missing data point when the lead concentration is below the detection limit. I use the lead concentration when the non-detected value is treated as 0 in my main analysis and use the other value for robustness checks.

The average lead concentration in Texas from 2006-2011 is 0.0019 mg/L with a standard deviation of 0.0011. Figure 3.5 shows the temporal and spatial variation in lead concentrations. As Panel (a) of Figure 3.5 shows, the lead concentration in Texas remains fairly constant over time with a small decrease in 2009. Panel (b) shows that lead concentration varies largely across counties. The counties with higher lead concentrations are mostly in central and eastern Texas, corresponding to the state's population distribution.



(a) By year (b) By county

Figure 3.3: Lead concentration by year and county 2006-2011

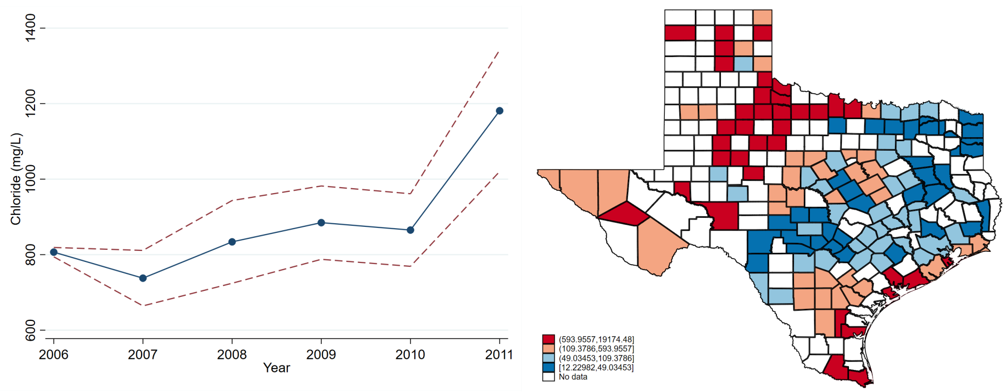
Notes: Panel (a) shows year fixed effects plus a constant from regressions that control for county fixed effects following Keiser & Shapiro (2019*c*). Blue connected dots show yearly values and red dashed lines show the 95% confidence interval for year fixed effect estimates, with 2006 as the reference year. Standard errors are clustered by county. Panel (b) shows the county average lead concentration over the full period, 2006-2011.

3.5.3 Surface water quality data

Surface water quality data come from Kuwayama et al. (2020). Chloride (Cl^{-}) concentration in surface waters was collected from multiple publicly available data sources. The first data source is the Water Quality Portal, a platform that provides water quality data from the USGS National Water Information System, the EPA STOrage and RETrieval Data Warehouse (STORET), and the U.S. Department of Agriculture's Sustaining The Earth's Watershed-Agricultural Research Database System. Additional Cl^{-} data are collected from the Surface Water Quality Monitoring Information System provided by TCEQ.

Surface water quality observations are obtained at the monitoring station level. I obtain the watershed boundary database from the Texas Natural Resources Information System (TNRIS), which is derived from the 1:24,000 USGS National Hydrography Dataset (NHD) (Texas Natural Resources Information System 2014). I define the Cl^- level of a watershed in a given year by calculating the average Cl^- concentration of all water quality monitoring stations located within the watershed using ArcGIS. I also obtain the county boundary shapefile from the US Census Bureau (US Census Bureau 2020) and use ArcGIS to estimate the area of overlap between a county and a watershed. I then estimate the average concentration of Cl^- of a county in a given year weighted by the area of overlap.

There is large variation in the Cl^- concentration across counties and over time in Texas. The weighted average Cl^- level is 190.7 mg/L in Texas surface water with a standard deviation of 473.4 mg/L. Figure 3.6 shows the temporal and spatial variation of surface water Cl^- concentration data in Texas. Figure 3.6a shows the year fixed effect plus a constant from regressions that control for county fixed effects following Keiser & Shapiro (2019*c*). As it is shown in Figure 3.6a, the average Cl^- concentration increases from 2006 to 2011. 3.6b shows the spatial distribution of surface water Cl^- concentration by county. Cl^- concentrations tend to be higher in northwestern and southeastern Texas.



(a) By year (b) By county

Figure 3.4: Chloride concentration by year and county 2006-2011

Notes: Panel (a) shows year fixed effects plus a constant from regressions that control for county fixed effects following Keiser & Shapiro (2019*c*). Blue connected dots show yearly values and red dashed lines show the 95% confidence interval for year fixed effect estimates, with 2006 as the reference year. Standard errors are clustered by county. Panel (b) shows the county average chloride concentration over the full period, 2006-2011.

3.5.4 Lead pipes data

The presence of lead pipes and cities' historical water characteristics data are from Clay et al. (2014*a*). The authors collect information on cities' use of lead pipes from the Manual of American Water-Works (Baker 1897). Clay et al. (2014*a*) collect data on 172 large and medium U.S. cities in 1900. When I match city-level lead data to the county in which each city is located, only have 5 counties in Texas have information on lead pipes' presence in 1900: Galveston, Harris, Jefferson, McLennan, and Tarrant counties. Figure 3.7 shows in location these counties within the state. Those without lead pipes present in 1900 as shaded yellow and those with lead pipes are shaded pink. I also map the top 11 big cities by population in Texas in 2019 with red circles. As it is shown in the figure, the presence of historical lead pipes is not strongly correlated with the current level of urbanization.

Following Clay et al. (2014a), I use two lead pipe variables. One variable is a categorical indicator of the presence of lead pipes. It is coded as 1 if the city has pipes made only with lead or a mix of lead and non-lead service pipes. Table 3.2 shows that 2% of the younger sample in my study are born in counties with the historical presence of lead pipes.

The other variable is the interaction of lead pipes and surface water characteristics. Because water chemistry affects corrosivity, thus affecting the lead concentration in drinking water, the authors also create a categorical variable using the interaction of lead pipes with the pH in source water [4]. I use the first lead pipes variable in the main analysis and use the second variable as a robustness check.

3.5.5 Additional controls

Neighborhood characteristics controls include the country year unemployment rate, median home income, poverty rate, and population. Unemployment and population data come from the Bureau of Labor Statistic's Local Area Unemployment Statistics (LAUS) county data. Income and poverty

[4]The presence of lead pipes is coded as having no lead pipes, having a mix of leaded pipes and no lead pipes, and only lead pipes. The pH in source water is coded as if the pH is below or above 7.3 (the median pH level in the data)

estimates come from the Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program county estimates.

Summary statistics from Table 3.2 show that the average unemployment rates for the younger cohorts and older cohorts are about 6.5% and 7.3% respectively. The average poverty rate is 17.9% for the younger cohort and 19.5% for the older cohorts. The median household income for the younger cohorts is $48757 on average and the median income for the older cohort is $31427. The average county population is 1591577 for the younger cohorts and 1011788 for the older cohorts.

I also include a list of weather control variables as they may affect educational outcomes through their impact on drinking water consumption. There is also evidence that temperature change over time affects the biochemical attributes of rivers (Ouellet et al. 2020). Weather controls data are from Schlenker & Roberts (2009), obtained from the author's website, which has daily minimum and maximum temperature and total precipitation on a 2.5×2.5 mile grid for the U.S. from 1900 to 2019. I estimate county-year average maximum temperature and precipitation in Texas counties using these daily variables.

3.6 Results

3.6.1 OLS results

Table 3.3 shows OLS estimates from equation (3.1), where Panel A presents results using the scaled standardized test scores and Panel B presents results where the outcome variable is an indicator of whether or not an individual meets the standard. I add control variables from left to right. Columns 1 to 5 show the OLS estimates of the impact of lead concentration on reading scores and Columns 6 to 10 show the estimated impacts on math scores. Columns 1 and 6 include only birth county and birth year fixed effects. Columns 2 and 7 show the results of adding individual controls, including gender, race, and economic disadvantage status. Columns 3 and 8 add controls for additional neighborhood characteristics, such as median household income, unemployment rate, and the poverty rate of the birth county in the birth year. I add county-by-year level average maximum temperature and precipitation in Columns 4 and 9. Columns 5 and 10 are robustness checks by including birth county by birth year trends.

Table 3.3: OLS estimates of the impacts of lead concentration on 3rd grade standardized test scores

A	3rd grade reading					3rd grade math				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(lead)	-2.335***	-2.296***	-2.191***	-2.029***	-1.299**	-1.2	-1.235*	-0.882	-0.863	0.898
	(0.624)	(0.609)	(0.696)	(0.684)	(0.506)	(0.758)	(0.706)	(0.817)	(0.774)	(0.976)
Observations	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142
R2	0.021	0.086	0.086	0.086	0.086	0.028	0.088	0.088	0.088	0.089
B	Met reading standard					Met math standard				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(lead)	-0.00636***	-0.0062***	-0.00606***	-0.00555***	-0.00334**	-0.00301	-0.00302*	-0.00186	-0.00199	0.00361*
	(0.00202)	(0.00209)	(0.00202)	(0.00208)	(0.00148)	90.00191)	(0.00175)	(0.00183)	(0.00168)	(0.00194)
N	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142	1323142
R2	0.018	0.093	0.093	0.093	0.093	0.021	0.084	0.084	0.084	0.084
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Standard errors in parentheses and clustered at county level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1-5 in Panel A of Table 3.3 suggest that the lead concentration in the year of birth reduces 3rd grade reading scores. The estimated coefficients on log(lead) are fairly consistent as more control variables are added [5]. Since I do not have controls that have been found highly predictive of child outcomes, such as maternal education, birth order, birth weight, or maternal education, the consistent magnitude of coefficients shows similar results with the past literature that a set of basic control variables may be sufficient to control for confounding bias (Aizer & Currie 2019). The coefficient in Column 1, -2.335, implies that a 1% increase in drinking water lead concentration decreases the 3rd-grade standardized reading scores by 0.023 points. This change seems small. But recall that the mean lead concentration in Texas (1.856 ppb) is well below the action level (15 ppb). Even so, if the average lead concentration in Texas increased to the 15 ppb action level under the LCR, the mean reading scores in Texas would decrease by 16.7 points (719 times 0.023), or 1.16 percent of the average reading score. Past studies have found that a mean BLL increase by 10 ppb is associated with a decline of average 3rd and 4th grade English Language Arts (ELA) scores by 2.6 percentage points in Massachusetts (Reyes 2015*b*) and a change of 3rd-grade reading scores by -0.335 points in Rhode Island children(Aizer et al. 2018). Because my study uses lead concentration in drinking water, not BLL, it is hard to compare these two measures. However, the fact that I find a statistically significant negative impact at levels two orders of magnitude smaller than the federal action level is notable.

Columns 6-10 in Panel A report suggest mixed results for the impact of

[5] I also checked the asinh transformation of lead as a robustness check and it did not change the main results (Table available upon request)

birth-year exposure to lead in drinking water on 3rd grade math scores. The estimated lead coefficients are of the expected signs but only weakly significant when I include birth county FE, birth year FE and individual controls, and insignificant in the other specifications. Estimated coefficients are also smaller than those in the reading models. If we interpret the insignificant Column 6 coefficient, a one percent increase in the drinking water lead concentration is associated with a decrease in 3rd-grade math scores of 0.012 points. The smaller impact on math scores is consistent with the epidemiological literature that suggests lead has a stronger relationship with verbal functioning (Bellinger et al. 1992, CDC 2004).

Panel B of Table 3.3 shows the regression results using Equation (3.1) where the dependent variable is an indicator variable set equal to one for individuals that meet STAAR standards for reading and math. Because the dependent variable is an indicator, I use linear probability models. Results are consistent with those from Panel A. Drinking water lead exposure in the birth year reduces the probability of passing reading tests,robust to controlling for various confounding factors. Impacts on the likelihood of passing the 3rd grade math test are not robust, and there is one counter-intuitive, positive and weakly significant coefficient in Column 10.

3.6.2 Instrumental variables results

Table 3.4 shows the first stage results from using Cl^- and the interaction between Cl^- and the historical presence of lead pipes as instruments for a lead action level exceedance (ALE). Columns 1-5 are estimated using equation (3.2), where I add control variables as one moves across columns. The coefficients on the mean Cl^- level are all positive and significant, suggesting an association between surface water Cl^- concentration and drinking water lead ALE that is robust to the addition of individual and neighborhood controls. The coefficient on mean Cl^- level in Column 4 suggests that holding individual characteristics, neighborhood characteristics and weather constant, the increase in the probability of having an ALE in one's birth county-year from a one-unit increase in surface water Cl^- concentration is 0.0184 percentage points.

The first stage estimates using the second instrument are in Columns 6-10 of Table 3.4. Similar to the first stage results for the first instrument, the interaction of source water Cl^- and historical lead pipes also has a positive and significant relationship with the drinking water lead concentration. The coefficient in Column 9 suggests that holding individual and neighborhood characteristics constant, in cities with lead pipes in 1900, the lead concentration in drinking water increases by 0.00148 ppb (0.08% of the average lead concentration) for each 1 mg/L (0.5% of average Cl^- level) rise in the source water Cl^- concentration.

I also report the first stage F statistics and the Sanderson-Windmeijer (SW) first-stage under-identification test in Table 3.4. The F statistics are lower than 10 for the first IV, suggesting the first instrument could be a weak instrument. The F statistics for the second IV is well above 10 suggesting the

Table 3.4: First stage result of IV strategies

	ALE					Lead concentration				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mean chloride	0.000209**	0.000209**	0.000191**	0.000184**	0.000188**					
	0.0000845)	(0.0000845)	(0.0000782)	(0.0000777)	(0.0000789)					
Mean chloride * presence of lead pipes in 1900						0.00165***	0.00165***	0.00155***	0.00148***	0.00143***
						(0.0000270)	(0.0000266)	(0.0000369)	(0.0000644)	(0.000145)
Observations	1318783	1318783	1318783	1318783	1318783	361103	361104	361105	361106	361106
R2	0.347	0.347	0.405	0.417	0.408	0.744	0.744	0.865	0.944	0.946
F statistics	6.12	6.12	5.97	5.67	6.79	3748.53	3860.22	1761.13	529.17	30.86
SW Chi-sq test (p-value)	0.013	0.0131	0.0142	0.0169	0.009	0.000	0.000	0.000	0.000	0.000
Hansens J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Standard errors in parentheses and clustered at county level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports the first-stage regression results of the two instrumental variables. Columns 1 to 5 report results on ALE and Columns 6 through 10 report the results on lead concentration. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include county-by-year trends. Standard errors in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

second instrument is unlikely to be a weak instrument(Stock & Yogo 2005). The SW chi-square under-identification tests suggest that both instruments are relevant. The over-identification test of both instruments (Hansens J statistics) is 0, suggesting both models are just identified.

Panel A of Table 3.5 reports second-stage results from the Cl^{-}-only IV models. Since the F statistics from the first stage are weak, I use the LIML estimator instead of 2SLS since it is more robust to weak instruments (Stock et al. 2002). I partial out the fixed effects and controls since the covariance matrix of orthogonality conditions S is not of full rank, and the overidentification tests are infeasible when I'm clustering standard errors (Baum et al. 2007). Columns 1 to 5 present the effects of ALE on 3rd grade scaled reading scores and Columns 6 to 10 present the effects on scaled math scores. The coefficient in Column 4 suggests that after controlling for individual, neighborhood, and weather characteristics, having a lead concentration in drinking water over 15 ppb is associated with an 18.3-point decrease in 3rd-grade reading scores.

Table 3.6 presents regression results using the second IV. The coefficients I present here are 2SLS estimates. LIML estimators of the approach are identical and available upon request. Panel A of Table 3.6 shows estimates of lead impact on 3rd-grade scores with additional controls added across the columns. In all the specifications, the average lead concentration has a negative and statistically significant impact on 3rd-grade standardized test scores in both subjects, except for Column 4. Contrary to previous findings, the impact of lead seems to be larger for math scores than reading scores. Column 5 reports the effect of lead on 3rd-grade reading scores controlling for individual, neighborhood characteristics, as well as a birth county by year trends. The coefficient suggests that a 1 ppb increase in average lead concentration is associated with a 3.911 point decrease in 3rd-grade reading scores. If the average lead concentration in drinking water increased to 15 ppb, a 13 ppb raise from the current average lead concentration in Texas, the associated decrease in 3rd-grade reading scores would be 50.8 points, 3.6 percent of the average reading scores. The magnitudes of the IV estimates are similar to the OLS

Table 3.5: IV estimates of lead impact on 3rd grade test scores using chloride level as instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	Reading scores					Math scores				
ALE	-17.99***	-19.83***	-18.99***	-18.3***	-14.88**	-4.116	-5.480	-4.044	-1.525	5.065
	(3.257)	(5.126)	(5.583)	(5.17)	(6.856)	(3.725)	(5.035)	(5.391)	(6.804)	(6.513)
B	Meet reading standard					Meet math standard				
ALE	-0.647***	-0.0701***	-0.0682***	-0.0676***	-0.0646*	-0.00791	-0.0117	-0.0059	0.00458	0.0582
	(0.00853)	(0.0129)	(0.0143)	(0.0131)	(0.0355)	(0.0144)	(0.0149)	(0.0166)	(0.0181)	(0.046)
Observations	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783
First stage F statistics	6.12	6.12	5.97	5.67	6.79	6.12	6.12	5.97	5.67	6.79
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table reports the effects of lead exposure at birth on 3rd-grade standardized test scores using instrumental variables. Panel A reports the effects of having an ALE at birth on 3rd-grade reading and math scores using the surface water Cl^- as an instrument. Panel B reports the effects of average lead concentration on 3rd-grade reading and math test results using the interaction of surface water Cl^- and historical presence of lead pipes as an instrument. Columns 1 to 5 report results on reading and Columns 6 through 10 report the results on math. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

estimates. Column 10 shows the equivalent estimates of lead impact on 3rd-grade math scores. A 1-ppb increase in the drinking water lead concentration is associated with a 5.198 point decrease in math scores. If the average lead concentration increase tod 15 ppb, the average 3rd-grade math scores would decrease by 4.6 percentage points.

Panel B of Table 3.6 presents IV results using the indicator for whether

Table 3.6: IV estimates of lead impact on 3rd grade test scores using interaction of chloride and lead pipes as instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	Reading scores					Math scores				
Lead	-0.910***	-2.192***	-2.304***	-1.453	-3.911**	-2.416***	-3.537***	-4.093***	-5.637***	-5.198**
	(0.061)	(0.039)	(0.121)	(0.886)	(1.956)	(0.0952)	(0.0403)	(0.0625)	(0.520)	(2.345)
B	Meet reading standard					Meet math standard				
Lead	-0.00119***	-0.00494***	-0.00605***	-0.00844***	-0.00217	-0.00675***	-0.00999***	-0.0112***	-0.0174***	-0.00645
	(0.000273)	(0.0000834)	(0.000502)	(0.000883)	(0.00604)	(0.000181)	(0.00005)	(0.00016)	(0.0011)	90.00438)
Observations	361106	361106	361106	361106	361106	361106	361106	361106	361106	361106
First stage F statistics	3748.53	3860.22	1761.13	529.17	30.86	3748.53	3860.22	1761.13	529.17	30.86
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table reports the effects of average lead concentration on 3rd-grade reading and math test results using the interaction of surface water Cl^- and the historical presence of lead pipes as an instrument. Columns 1 to 5 report results on reading and Columns 6 through 10 report the results on math. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

an individual meets the reading and math standards, using the second instrument (Cl^{-} $\times$ lead pipes) . Since these dependent variables are dummy variables, the models reported here are linear probability models. Following the previous tables, I report results for reading standards in Columns 1 to 5 and results for math standards in Columns 6 to 10. Coefficients in Panel B suggest that the drinking water lead concentration at birth has robust, significant negative impacts on students' probability of passing 3rd grade reading and math standards. The coefficient in Column 4 of Panel B suggests that a 1 ppb increase in average lead concentration leads to a decrease in the probability of meeting the 3rd-grade reading standard by 0.8 percent, holding all else constant. From Column 8, a 1 ppb increase in lead concentration is associated with a 1.74% decrease in the probability of passing the math standardized test. The coefficients on the lead are negative but insignificant in Columns 5 and 10. This may be caused by the fact that I only have 5 counties with the lead pipes data so the birth county by year trend absorbs a lot of variation.

Given the robustness of the results, the IV strategy using the interaction of surface Cl^{-} concentration and presence of lead pipes may be the preferred approach to estimate the impact of lead in drinking water. However, it is important to note that the sample size is significantly smaller than the OLS sample because I must have data on the presence of lead pipes in 1900. In Texas, this information is available for only 5 counties. Even though I find consistent negative impacts of lead in drinking water on academic performance, the IV sample may not be representative of the impact of lead statewide. For these reasons, I use the fixed effect model to examine the long-run effect of lead in the next section and plan to apply this approach to the whole U.S. in the future.

3.6.3 Long-run impacts of LCR violations

The results presented so far are based on the availability of lead concentration data and the IV. However, given that I only have lead concentration data from 2006-2011, I cannot estimate the long-run impacts of lead exposure with these strategies because children born in 2006-2011 have not yet graduated from high school. Since the LCR violations have plausibly exogenous timing, I use a difference-in-difference design to exploit changes in LCR violation status across counties and examine this long-run question. With the difference-in-difference design, I compare the difference in outcomes between cohorts born in counties with LCR violations and cohorts born in counties without LCR violations.

The key assumption underlying this identification strategy is that treatment and control counties would have the same trends in high school graduation rates without the LCR violation. To test the validation of the parallel trends assumption, I regress the high school graduation on the interaction of LCR violation status at birth and the years since the LCR violation initially occurred, including birth county and birth year fixed effects. Figure 3.8 shows the coefficient for the interaction term with 1 year before the LCR violation as the reference year. In the pre-period, the coefficients of LCR violation are small and insignificant, suggesting the parallel trends assumption holds. The treatment is negative and statistically significant in periods T+1 and T+2 suggesting that an LCR has negative impacts on high school graduation rate persistent for 2 years. It is important to note that the coefficients of years T-8 to T-6 and T+4 to T+6 have large standard errors. This is because the majority of LCR violations happen in 1992 and last for an average 3 years, so the coefficients for the years before T-6 and after T+4 are less precisely estimated.

I estimate the DID model using the cohort born between 1990 and 2001 enrolled in a Texas public school in 9th grade. Table 3.7 presents the regression results with Equation (3.4). Panel A shows results from estimating the impacts of lead in drinking water on high school graduation, and Panel B shows results for public university enrollment in Texas.

Results in Panel A suggest that an LCR violation in the birth year has a significant robust negative impact on the high school graduation rate. Column 4 suggests that holding all else constant, students experiencing an LCR violation at birth are 0.6 percent less likely to graduate high school. Panel B results suggest that, contrary to the high school graduation rate, an LCR violation at birth does not have an impact on public university enrollment in Texas. It is also important to note that the ERC data only has information on an individual's enrollment in Texas universities. A child who enrolls in a private university in Texas or any university outside of the state does not show up in these data. Moreover, Texas has the Top Ten Percent rule where anyone

Table 3.7: Long run impact on high school graduation rate and public university enrollment

	(1)	(2)	(3)	(4)	(5)
			A: High school graduation rate		
LCR	-0.00860***	-0.00861***	-0.00571***	-0.00574***	-0.00526***
	(0.00320)	(0.00286)	(0.00263)	(0.00263)	(0.00300)
Observations	1341729	1341729	1341729	1341729	1341729
R2	0.024	0.052	0.052	0.052	0.053
			B: Public university enrollment		
LCR	-0.00236	-0.00187	0.00242	0.00267	0.00353*
	(-0.00278)	(-0.00271)	(-0.00222)	(-0.00232)	(-0.00192)
Observations	1341757	1341757	1341757	1341757	1341757
R2	0.023	0.088	0.088	0.088	0.088
			C: High school graduation rate with all violation years		
LCR	-0.00758***	-0.00671***	-0.00252	-0.00257	-0.00398
	(0.00257)	(0.00250)	(0.00190)	(0.00193)	(0.00270)
Observations	1341757	1341757	1341757	1341757	1341757
R2	0.024	0.052	0.052	0.052	0.053
Birth year FE	Yes	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes

Notes: Panel A reports the effects of LCR violation exposure at birth on high school graduation rate and Panel B reports the effects of LCR violation exposure at birth on Texas public university enrollment rate. Panel C reports the effects of LCR violation exposure at birth on high school graduation rate using all violation years. Column 1 reports the coefficient estimates using only birth county and birth year FEs. Column 2 includes individual-level controls. Column 3 includes neighborhood controls including poverty rate, unemployment rate, and median household income. Column 4 includes weather control variables. Column 5 includes birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

in the top 10 percent of their high school class is guaranteed entrance to a public university. This rule may also attenuate my coefficient on public university enrollment. My estimate may be an underestimate of the true effect of early exposure to drinking water lead treatment violation on college enrollment.

It is important to note that I only include those who were born with a new LCR violation in their birth county in my main DID sample. Given that the average duration of an LCR violation in Texas is about 3 years, and PWSs are required to notify the public about the violation, parents of individuals who were born in the 2nd or 3rd year of an LCR violation may be aware of the potentially high level of lead in their drinking water and take avoidance actions. In Panel C of Table 3.7, I use the sample of people who are born with any year of an LCR violation. The coefficients in Panel C are smaller and less robust, consistent with potential mitigation methods may be taken for the cohorts born with existing LCR violations in their birth counties.

3.6.4 Heterogeneity in impacts of lead exposure

I examine the heterogeneous effects of lead exposure by individual demographics, such as gender and race, using regressions for the second IV strategy and the DID strategy. I include an interaction term of my key treatment variables with indicators for race, gender, and economic status. Since the models contain interaction terms between instrument/treatment and the group-specific dummy variable, I also include the group dummy variables in the regressions, so the effect on each group is the sum of these two coefficients. Table 3.8 presents results for 3rd-grade standardized test scores (Panel A and B) and high school graduation rate (Panel C). Columns 1 and 2 show the heterogeneous impact of lead exposure by gender. As is shown in Columns 1 and

2 of Panel A and Panel B, the interaction coefficient is significantly positive for male students and significantly negative for female students, suggesting that female students' 3rd-grade standardized test scores may be more vulnerable to lead exposure.

Columns 3 through 7 of Table 3.8 suggest that there may be significant heterogeneous treatment effects of lead exposure at birth on later educational outcomes across different racial groups. Black students are disproportionately impacted by lead exposure at birth on both 3rd-grade reading and math scores and also on high school graduation rates. Surprisingly, white students in Texas experience more negative impacts from drinking water lead exposure at 3rd grade, but less negative impacts when they are graduating from high school. Asian students and Hispanic students also experience more negative impacts from drinking water lead exposure by 3rd grade.

Columns 8 to 10 of Table 3.8 present the heterogeneous effects of lead by economic status. Children born to families with economic disadvantages may be disproportionately affected by drinking water lead exposure at birth.

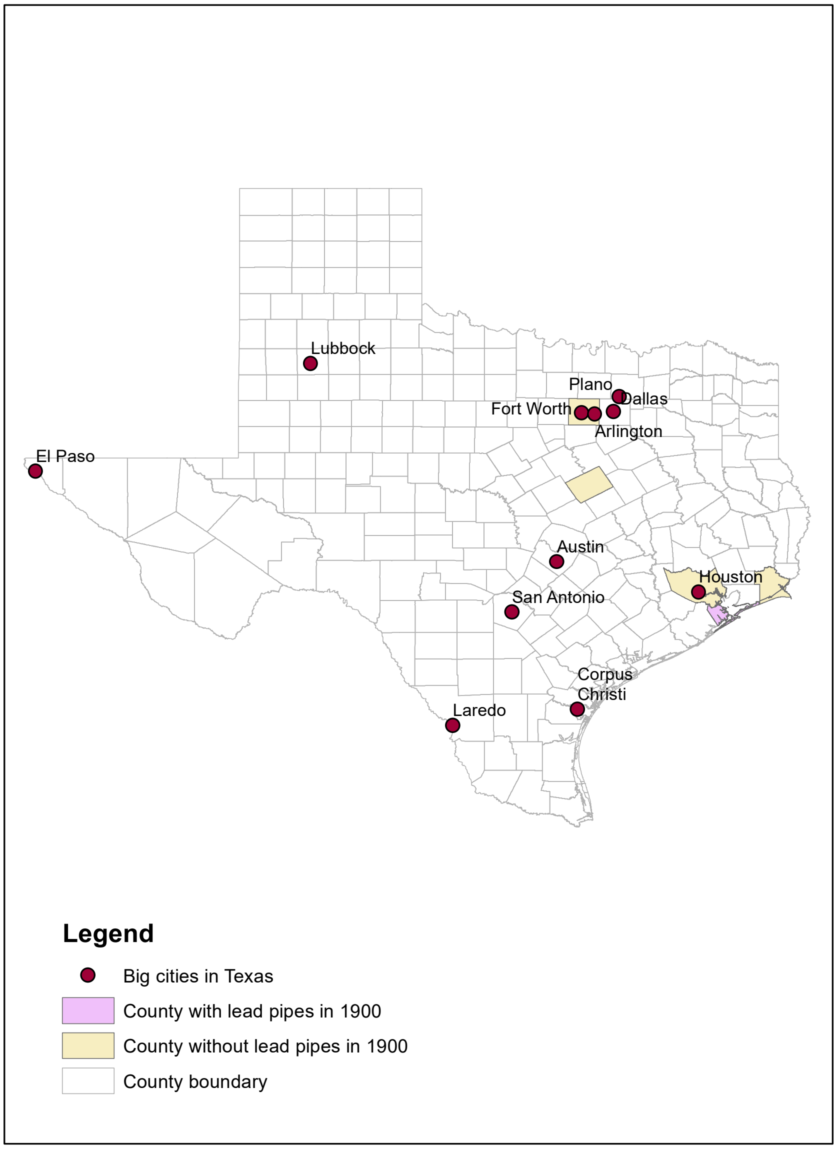


Figure 3.5: Texas counties with information on lead pipes in 1900.

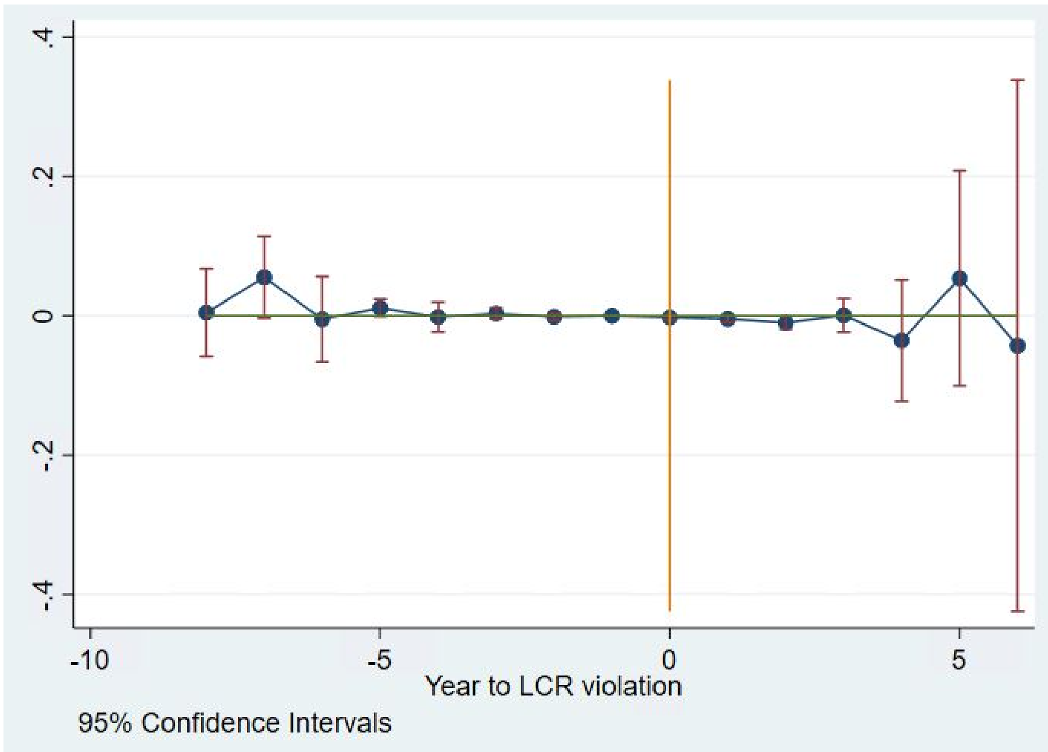


Figure 3.6: The estimated treatment effects by event time

Notes: Figure presents estimated trends in high school graduation rate using sample from the primary analyses plus years "T-8" to "T+6" to better map out the pre-treatment periods and treatment response. I regress the high school graduation rate with an interaction of LCR violation status and time dummies for all period before and after treatment. Blue connected dots show coefficient and red dashed lines show the 95% confidence interval, with the year before an LCR violation as the reference year. I calculate robust, individual level standard errors.

Table 3.8: The effect of lead exposure on 3rd grade scores and high school graduation rate by gender, race and economic status

	(1) Male	(2) Female	(3) White	(4) African American	(5) Asian	(6) Hispanic	(7) Native American	(8) Free lunch	(9) Reduced lunch	(10) Other disadvantages
Panel A:3rd grade reading										
Second instrument	-0.003***	0.001*	0.003**	0.001*	0.00009	-0.0004	-0.0003	0.006***	-0.0001	-0.001*
	(0.0003)	(0.0006)	(0.0004)	(0.0005)	(0.0005)	(0.0008)	(0.0006)	(0.0004)	(0.0003)	(0.0003)
Second instrument * group	0.005***	-0.005***	-0.010***	-0.014**	-0.004***	0.0008	0.008**	-0.014***	-0.006***	0.010***
	(0.0003)	(0.0003)	(0.001)	(0.003)	(0.0003)	(0.001)	(0.001)	(0.001)	(0.0001)	(0.001)
R2	0.0814	0.0814	0.0689	0.0678	0.07	0.0647	0.0633	0.073	0.0621	0.063
Panel B:3rd grade math										
Second instrument	-0.010***	-0.006***	-0.004*	-0.007***	-0.007***	-0.006***	-0.007***	-0.004***	-0.007***	-0.008***
	(0.0005)	(0.0007)	(0.001)	(0.0004)	(0.0005)	(0.0008)	(0.0006)	(0.0003)	(0.0004)	(0.0004)
Second instrument * group	0.004***	-0.004***	-0.009**	-0.008*	-0.008***	-0.004**	0.018**	-0.009***	-0.011***	-0.0087***
	(0.0005)	(0.0005)	(0.002)	(0.002)	(0.0002)	(0.001)	(0.004)	(0.001)	(0.0002)	(0.001)
R2	0.0851	0.0851	0.0592	0.0696	0.0685	0.0554	0.0554	0.0794	0.071	0.0714
N	361024	361024	361024	361024	361024	361024	361024	361024	361024	361024
Panel C:High school graduate										
LCR	-0.037*	-0.040**	-0.058***	-0.021	-0.039***	-0.031**	-0.038***	-0.033**	-0.037***	-0.034**
	(0.020)	(0.025)	(0.015)	(0.015)	(0.014)	(0.015)	(0.014)	(0.016)	(0.014)	(0.014)
LCR * group	-0.003	0.003	0.063***	-0.074***	0.059	-0.018	-0.054	-0.010	-0.018	-0.070
	(0.026)	(0.015)	(0.019)	(0.016)	(0.052)	(0.017)	(0.113)	(0.016)	(0.019)	(0.043)
R2	0.050	0.050	0.047	0.047	0.049	0.047	0.047	0.047	0.047	0.047
N	1341714	1341714	1341714	1341714	1341741	1341741	1341741	1341741	1341741	1341741
Birth year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

This table reports regression coefficients from 30 separate regressions. The regressions are estimated using the individual level data of the two samples. The dependent variables are 3rd grade reading scores, math scores and high school graduation rate. Panel A and Panel B use the IV sample and Panel C uses the DID sample. Panel A and Panel B reports estimates from reduced-form models that include additional interaction terms between the second instrument and a group-specific dummy variable indicated in the column heading. I also the group-specific dummy variable and the instrument variables separately in the regression models. Panel C reports estimates from the DID model with the additional interaction term of the LCR exposure at birth and the group dummy, a group-specific dummy and LCR violation status at birth. I also include birth year FE, birth county FE, individual controls, neighborhood controls and weather controls as described before. Standard errors in parentheses and clustered at county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.5 Robustness checks

In this section, I explore the robustness of my results to a variety of additional tests and specifications.

3.6.5.1 Include median house age as control

A common pathway of lead entering into children's bodies is through lead paint and the contaminated dust and soils it generates (Lanphear & Roghmann 1997, Jacobs et al. 2002). Though lead paint was banned in 1978, older homes are more likely to still have lead paint and lead pipes. To control for the potential presence of lead paint, I include the median home age in the birth county in the birth year as an additional control variable in the first IV strategy and report the coefficients in Table 3.9.

Column 1 of Table 3.9 reports the first-stage result using the interaction of surface water Cl^- concentration and historical presence of lead pipes as an instrument. Columns 2 and 3 show results for 3rd-grade reading and Columns 4 and 5 report results for math. All coefficients are estimated with birth year FE, birth county FE, individual controls, neighborhood controls, and weather controls as before.

Columns 2 and 4 are consistent with previous results that having lead exposure at birth significantly reduces 3rd-grade reading and math scores. Relative to the baseline results without median home age controls, the effect of drinking water lead on 3rd-grade scores is slightly larger.

3.6.5.2 Include county trends

Tables 3.5, 3.6 and 3.7 also presents additional results including county by year trends, to address the concern that broad trends at the county level might be influencing my results. The results are of the same sign and similar magnitude to the main results.

Table 3.9: Robustness checks with housing age

	Lead	Reading		Math	
	(1)	(2) Scores	(3) Meeting standards	(4) Scores	(5) Meeting standards
Chloride * lead pipe	0.00148***				
	(0.0000644)				
Lead		-2.647*	-0.00944***	-8.844***	-0.0271***
		(1.365)	(0.00107)	(0.939)	(0.00205)
Observations	361106	361106	361106	361106	361106
First stage F statistics	134.64	134.64	134.64	134.64	134.64
SW Chi-sq (p-value)	0.000				
Hansens J statistics	0.000				

Notes: This table reports regression coefficients using the interaction of surface water Cl^- and the historical presence of lead pipes as an instrument. Column 1 reports first stage results. Columns 2 and 3 show results on 3rd-grade reading results and Columns 4 and 5 are results on math. Columns 2 and 4 show results on test scores and Columns 3 and 5 show results on indicators of whether a student passes the tests. I also include birth year FE, birth county FE, individual controls, neighborhood controls, and weather controls as described before. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.5.3 Control for PWS size

As noted before, LCR violations are more likely to happen in small PWSs. People born in households served by small PWSs may differ from people born in larger PWSs in unobservable ways. As a robustness check, I estimate the second IV model using the interaction between Cl^{-} concentration and lead pipes and controlling for the sum of the population served by PWSs and the number of small and very small PWSs in a county. Table 3.10 reports the results, where Columns 1 and 2 control for the population served by PWSs, and Columns 3 and 4 control for the number of small and very small systems. Columns 1 and 3 are results from specification 3.4, while Columns 2 and 4 add birth county by birth year trends. Comparing the coefficient estimates in Table 3.10 to those in Table 3.7, controlling for PWSs sizes has little effect on the educational impact of LCR violations at birth.

Table 3.10: Robustness checks controlling for PWS size

	(1)	(2)	(3)	(4)
	Population served		Number of small and very small systems	
LCR	-0.00575**	-0.00527*	-0.00575**	-0.00553*
	(0.00263)	(0.003)	(0.00263)	(0.00305)
Observations	1340066	1340066	1340066	1340066
R2	0.052	0.053	0.052	0.053
Birth year FE	Yes	Yes	Yes	Yes
Birth county FE	Yes	Yes	Yes	Yes
County by year trend	No	Yes	No	Yes
Individual controls	No	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes
Weather controls	No	No	No	Yes

Notes: This table reports regression coefficients using the interaction of surface water Cl^- and the historical presence of lead pipes as an instrument. Columns 1 and 2 are results controlling for the population served by PWSs whereas Columns 3 and 4 are results controlling for the number of small and very small systems. Columns 1 and 3 are results controlling for birth year FE, birth county FE, individual controls, neighborhood controls, and weather controls as described before. Columns 2 and 4 add birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6.5.4 Use Chloride-Sulfate Mass Ratio (CSMR)

The last robustness check uses the Chloride-Sulfate Mass Ratio (CSMR) instead of Cl^- as an instrument. While high Cl^- can promote solubility of lead in drinking water, sulfate may inhibit corrosion of lead-bearing materials both in isolation and in galvanic connections to copper(Edwards & Triantafyllidou 2007). Environmental engineering experiments and utilities' practical experience have shown that the chloride-to-sulfate mass ratio (CSMR) may also be an important indicator to control lead leaching to potable water (Edwards & Triantafyllidou 2007). In Tables 3.11 and 3.12, I report results from a set of IV models using this alternative instrument, interacted with the historical presence of lead pipes.

In the first stage, the CSMR has a consistently positive impact on ALE and lead concentration. The F statistics on CSMR suggest that CSMR alone, like Cl^{-} is a weak instrument, but the interaction of CSMR and lead pipes is a strong instrument.

Panel A and Panel B in Table 3.12 report 2nd stage results using the CSMR as an alternative instrument. Panel C and Panel D report results using the interaction of CSMR and the historical presence of lead pipes as an instrument. Columns 1 to 5 report results on reading and Columns 6 through 10 report the results on math. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include birth county by birth year trends. Similar to the results using Cl^{-} as an instrument, ALE has consistent negative impacts on 3rd-grade reading tests and the lead concentration has negative and significant impacts on both reading and math results.

Table 3.11: First stage of IV using CSMR

	ALE					Lead concentration				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CSMR	0.0615	0.0615	0.0610*	0.0581	0.0615*					
	(0.0402)	(0.0402)	(0.0367)	(0.0354)	(0.0368)					
CSMR* presence of lead pipes in 1900						0.564***	0.565***	0.765**	0.657***	0.586***
						(0.00556)	(0.00561)	(0.137)	(0.107)	(0.027)
Observations	1318783	1318783	1318783	1318783	1318783	361106	361106	361106	361106	361106
R2	0.281	0.281	0.354	0.369	0.361	0.574	0.575	0.775	0.872	0.895
F statistics	2.34	2.34	2.77	2.79	3.08	10289.35	10119.34	31.24	38.06	121.97
SW Chi-sq test (p-value)	0.125	0.125	0.095	0.094	0.078	0.000	0.000	0.000	0.000	0.000
Hansens J statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table reports first stage results using CSMR as instruments. Columns 1 to 5 show impacts of CSMR on ALE and Columns 6 through 10 show results of interaction terms on lead concentration. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.12: IV estimates results using CSMR and lead pipes as instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	Reading scores					Math scores				
ALE	-14.52*	-18.25*	-22.69*	-25.55**	-16.50	-5.431	-9.267	-10.53	-5.265	20.34
	(8.529)	(9.583)	(10.25)	(11.08)	(14.03)	(10.87)	(12.12)	(13.08)	(12.76)	(18.33)
B	Meet reading standard					Meet math standard				
ALE	-0.0559**	-0.0679**	-0.0695**	-0.0829***	-0.0646*	-0.0284	-0.0403	-0.0456	-0.0206	0.0582
	(0.0268)	(0.030)	(0.0310)	(0.0309)	(0.0355)	(0.0273)	(0.0322)	(0.0354)	(0.0310)	(0.0460)
Observations	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783	1318783
First stage F statistics	2.34	2.34	2.77	2.79	3.08	2.34	2.34	2.77	2.79	3.08
C	Reading scores					Math scores				
Mean lead	-2.396***	-4.081***	-4.070***	-3.902***	-7.387***	-2.943***	-4.402***	-5.049***	-7.623***	-8.184***
	(0.144)	(0.220)	(0.932)	(1.124)	(2.307)	(0.292)	(0.222)	(0.623)	(0.743)	(2.241)
D	Meet reading standard					Meet math standard				
Mean lead	-0.00372*	-0.00858*	-0.0108**	-0.0128**	0	-0.00666***	-0.0110***	-0.0132***	-0.0225***	0
	(0.0005000)	(0.000237)	(0.00236)	(0.00305)	(.)	(0.000931)	(0.000730)	(0.00162)	(0.00286)	(.)
Observations	361106	361106	361106	361106	361106	361106	361106	361106	361106	
First stage F statistics	10289.35	10119.34	31.24	38.06	121.97	10289.35	10119.34	31.24	38.06	121.97
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Notes: This table reports the effects of average lead concentration on 3rd-grade reading and math test results using two instrumental variable strategies.Panel A and Panel B report results using CSMR as an instrument. Panel C and Panel D report results using the interaction of CSMR and the historical presence of lead pipes as an instrument.Columns 1 to 5 report results on reading and Columns 6 through 10 report the results on math. Columns 1 and 6 report the coefficient estimates using only birth county and birth year FEs. Columns 2 and 7 include individual-level controls. Columns 3 and 8 include neighborhood controls including poverty rate, unemployment rate, and median household income. Columns 4 and 9 include weather control variables. Columns 5 and 10 also include birth county by birth year trends. Standard errors are in parentheses and clustered at the county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.7 Discussion and Conclusions

Using restrictive education data and drinking water data in Texas from 1991 to 2011, this study investigates the educational impacts of drinking water lead exposure in early childhood. I exploit variation in lead concentrations via changes in surface water Cl^- concentration, which affects treated drinking water chemistry, as well as the historical presence of lead pipes. I also exploit the plausibly exogenous timing of LCR violations by comparing the outcomes of children born with and without a new LCR violation in their birth county-year.

My results provide empirical evidence that drinking water lead exposure in the birth year has significant negative impacts on standardized test scores in the third grade, even when exposure is well below the federal action level for lead in drinking water under the SDWA. Specifically, I find that a 1-ppb increase in lead concentration in drinking water is associated with a 4-points and 5.2-point decrease in 3rd-grade reading and math scores, respectively. Given the average concentration of lead is about 2 ppb, eliminating lead in Texas could increase average reading scores by 52 points and math scores by 65 points, corresponding to 30 percent and 37 percent of a standard deviation. These are relatively modest gains given that the average scores of both tests are around 1400, but the magnitude is actually bigger than the estimates for BLLs in the existing literature (Aizer & Currie 2019). Eliminating lead in drinking water is also associated with a 0.8% and 1.74% increase in the probability of passing reading and math tests in 3rd grade. These estimates are also consistent with the existing literature using BLLs in both economics and epidemiology (Aizer & Currie 2019). Importantly, the effects that I estimated occur in a state and at a time when the lead concentration in drinking water is, on average, quite low.

This study also suggests that violation of the federal regulation for drinking water lead treatment in the birth year leads to a smaller probability of graduating from high school. Results suggest that being born in a county with a new LCR violation is associated with a 0.6% decrease in the probability of graduating from high school. A high school drop-out's weekly income is \$606, \$143 less than people with high school diplomas, in the U.S (U.S. Bureau of Labor Statistics 2019). Thus, an LCR violation at birth may be associated with a \$45 or 0.14% decrease in average annual income through its impact on high school graduation. For the older cohort in my sample, the benefits in terms of increased income from eliminating LCR violations would be around \$12 million annually in Texas alone. My estimates are smaller than the only existing study of the long-term impact of lead, which suggests reducing childhood BLL from 10 $\mu g/dL$ to 5 $\mu g/dL$ in the U.S. would have a benefit of around \$198 million annually (Grönqvist et al. 2020). However, it is important to note that my study uses LCR violations that are associated with a very low level of lead concentration. Given the 1.85 ppb in lead, concentration is only about 4% of the 5 $\mu g/dL$ reduction, eliminating LCR violations is associated with a sizeable economic benefit.

If the \$45 income loss persists over the life cycle, an individual's lifetime wage loss alone from having an LCR violation at birth is about \$900 [6]. However, given an individual's wage is unlikely to be constant through the life cycle, Carnevale et al. (2011) estimate that workers with a high school diploma on average earn \$331,000 more than people without during their working lives. Exposing to an LCR violation at birth, then the loss in the present discounted value of lifetime earnings is about \$2,000. Given 20% of individuals in my sample are exposed to new LCR violations at birth, the loss in lifetime earnings ranges from around \$241 million to \$535 million in Texas.

President Biden's The American Jobs Plan includes \$45 billion to replace lead pipes and service lines across the U.S. While I do not have an accurate estimate of how many people are exposed to LCR violations at birth during 1991-2004, one can generalize my estimates assuming that the rate of the population exposed in Texas is the same as the rest of the country. Given the average number of births in the U.S. is around 4 million, 56 million people were born during the period. If 20% of them (11.2 million) are exposed to a new LCR violation at birth, the annual loss of income is about \$504 million. The loss in lifetime earnings would be around \$10 billion to \$22 billion. A back-of-the-envelope calculation based on EPA's estimate of average replacement cost per line (\$4,700) and assumption of 6 to 10 million lead service lines across the country suggests the cost could range from \$28 billion to \$47 billion (Campbell & Wessel 2021). The benefit from an increase in lifetime earnings alone could not cover the cost of replacing all lead pipes in the U.S. But since

[6] I assume an individual can work from 18 to 78 and use a discount rate of 5%.

this is just one slice of potential benefits from eliminating lead from drinking water, the total benefit is likely much larger.

I also find that female children and children from African American families and families with economic disadvantages are more vulnerable to lead exposure, in line with the existing literature (Chetty et al. 2016, Heckman & Karapakula 2019, Grönqvist et al. 2020). The findings suggest that early childhood lead exposure may be on contributing factor to the racial-achievement gaps in the U.S.

This study using data in Texas provides the first empirical evidence that early childhood exposure to lead from drinking water well below action levels has a significant impact on educational outcomes in both the short and the long run. However, my IV sample of 5 Texas counties and the comparably lower lead concentration in Texas may not provide a comprehensive estimate of drinking water lead impact on the whole U.S. An important direction of my future research is to take the approach and apply it to a broader geographic area. One educational data set that I plan to use is the school district level standard test scores data from the Stanford educational archive (Sorensen et al. 2019). I have also recently obtained access to the restrictive federal Census survey data. The federal restrictive data will enable future work evaluate drinking water lead impacts with a broader geographic scope than I have examined in this paper.

Chapter 4

Water Pollution Control in Developing Countries: Policy Instruments and Empirical Evidence [1]

4.1 Introduction

Severe water pollution problems are widespread in developing countries, where many major river systems are highly impaired. For example, within the monitored areas of China's main river systems, only 28 percent have water suitable for drinking, and about one-third do not meet the country's lowest ambient water quality standards (standards that focus on water quality within lakes, rivers, streams and other raw water sources), which makes these rivers unsuitable even for irrigation (World Bank 2006). India's Ganga River, alone, receives point-source pollution comprising more than 1.3 billion liters of untreated domestic waste and 260 million liters of untreated industrial waste, which is in addition to agricultural and urban runoff (Dakkak 2018). In China, Brazil, India, and Indonesia (the four most populous developing countries), dissolved oxygen levels, which are an important indicator of healthy

[1] This chapter is published in *Olmstead, S. and Zheng, J., 2021. Water Pollution Control in Developing Countries: Policy Instruments and Empirical Evidence. Review of Environmental Economics and Policy, 15(2), pp.000-000.* Olmstead and I designed the study, performed the literature review, and wrote the manuscript together.

aquatic ecosystems, are well below average levels in industrialized countries (Greenstone & Jack 2015).

Ambient water pollution in developing countries harms human health (Ebenstein 2012, Do et al. 2018, Garg et al. 2018). Poor water quality may also reduce agricultural output (Hagerty 2018), educational outcomes (Zhang & Xu 2016), and labor productivity (Meeks 2017). Polluted surface water in developing countries also causes damages to recreational opportunities and other ecosystem services (Choe et al. 1996, Day & Mourato 2002, Beharry-Borg et al. 2010, Mishra 2017). These damages suggest that water pollution control policies in developing countries are likely to generate substantial benefits. The standard economic approach to cost-effectively reduce municipal, agricultural, and industrial water pollution in developing countries would be to use market-based policy instruments. Even in industrialized countries, however, market-based policy instruments have proven to be more difficult to apply effectively to water pollution than to air pollution (Fisher-vanden & Olmstead 2013).

This article examines the policy instruments that can be used to control ambient water pollution in developing countries and reviews the empirical evidence in the economics literature on the effectiveness of these policy options in practice. The article is organized as follows. In the next four sections, we discuss the categories of policy instruments that can be used to control ambient water pollution – prescriptive approaches, market-based approaches, voluntary approaches, and infrastructure investments, respectively – and the results of empirical assessments of their effectiveness. We focus primarily on developing countries, but mention specific policies and empirical evidence from industrialized countries where appropriate. Then we highlight challenges to the effective design, implementation, and evaluation of water pollution control policies in developing country settings. These include sparse and/or poor quality data, inadequate monitoring and enforcement, rent-seeking in regulatory settings, and inter-jurisdictional spillovers when regulation is decentralized. In the final section, we summarize our findings, discuss the gaps in the empirical literature on the impacts of water pollution policy as well as some priorities for future research, and present some concluding thoughts.

4.2 Prescriptive Policies

The most common approach to environmental regulation focuses on prescriptive policy instruments, sometimes called command-and-control instruments, which regulate the behavior or performance of individual factories, power plants, and other commercial and industrial facilities. For example, a technology standard requires firms to use a particular pollution abatement technology. A performance standard may impose a maximum allowable emissions rate, and thus allow polluters more flexibility in the choice of control technology. Most existing water quality regulations in industrialized countries use these approaches. Similarly, developing countries regulate water pollution primarily through the use of prescriptive standards. For example, India's 1985 National River Conservation Plan (NRCP) established a set of designated uses for surface waters and approaches for achieving levels of water quality appropriate to those designated uses (Greenstone & Hanna 2014). Although the NRCP requires the construction of sewage treatment plants and other capital investments to reduce water pollution, it does not provide a dedicated source of revenues to fund those investments. Greenstone and Hanna (2014) find that India's NRCP has not reduced water pollution concentrations in river segments covered by the Plan. They argue that this failure is due to low public demand for ambient water quality improvements and weak institutional support (e.g., from the Supreme Court and monitoring agencies) for the NRCP's goals. In China, the main policy on water pollution reduction is currently the Water Pollution Prevention and Action Plan, known as the "Water Ten Plan" (State Council 2015). The Water Ten Plan establishes targets for water pollution reduction and approaches for achieving them, including setting pollution reduction targets for small factories in polluting industries and shutting down those that fail to meet the targets. The Water Ten Plan also requires plants in ten major polluting industries to install specific abatement technologies. For instance, all pulp and paper factories in China are required to switch to either Elemental Chlorine Free or Total Chlorine Free bleaching technologies. Empirical estimates suggest that the Water Ten Plan may be associated with reductions in water pollution (Wang & Wei 2019), but causal estimates have not yet appeared in the literature.

4.3 Market-based Policies

Market-based policy instruments are decentralized, focusing on aggregate or market-level outcomes, such as total pollution levels or total emissions, rather than individual facilities. Examples of such instruments include pollution taxes and subsidies, tradable pollution permits, payments for ecosystem services, and information disclosure. In this section, we describe each of these market-based water pollution policies as well as relevant applications in industrialized and developing countries.

4.3.1 Pollution Taxes and Subsidies

The standard market-based economic approach is to tax negative externalities and subsidize positive externalities. Under a uniform tax, marginal abatement costs are equal across firms, thus generating the least-cost allocation of emissions reductions. However, in the case of water pollution, a uniform tax is often not efficient, because marginal damages usually depend on the location of the source of the pollution. This means that an efficient tax must also vary by source or at a more aggregate level, for example, dividing sources into "zones" that recognize the spatial heterogeneity in damages (Boyd 2003). The standard Pigouvian tax, set equal to the marginal damages at the efficient level of pollution, easily addresses pollution from point sources. Taxing non-point source pollution (e.g., diffuse runoff from farms and cities), which generally accounts for a large share of total water pollution, is more complicated, partly because effluent is not easily monitored or measured, and has thus prompted much discussion in the literature (Shortle & Horan 2001). Water pollution taxes are fairly uncommon in industrialized countries. Although France, Germany, and the Netherlands all have systems of water pollution taxes, with some dating to the 1970s (Boyd 2003), only the Dutch water pollution fee system has been found to have a statistically significant association with pollution reductions (Bressers 1988). In contrast, a few developing countries have experimented with more robust water pollution taxation, and some of these policies have been evaluated econometrically.

China's pollution levy system, established in the early 1980s, initially required industrial plants to pay a fee on the (one) pollutant that exceeded the applicable standard by the greatest amount. Since a 2003 reform, plants must pay levies on the three pollutants that exceed the standard by the greatest amount, and the levy rates have increased dramatically. Even before the 2003 reform, analyses showed that the pollution levy system reduced emissions (Jiang & McKibbin 2002, Wang & Wheeler 2003). Based on data on plant-level pollution expenditures, Wang (2002) finds that industries respond strongly to pollution charges but not to other regulatory approaches. In another empirical analysis, Ebenstein (2012) finds that doubling China's levy for wastewater dumping would avert 17,000 premature deaths from digestive cancers per year at a cost of about $500 million per year, which is a fraction of value-of-statistical life (VSL) estimates for China.

Since 2009, a separate environmental tax – the "Pay for Permit" policy – has been applied to chemical oxygen demand (COD) emissions from industrial sources in the Lake Tai Basin in Jiangsu province. This pilot program charges firms for every unit of pollution (a classic Pigouvian tax). Participating firms are required to purchase a permit from the local government for each unit of expected COD emissions, with penalties for violations. He & Zhang (2018) find that participating plants reduced emissions by about 40 percent in the first two years of the policy.

Colombia implemented a national discharge fee system (Law 99) in 1993. The law mandated that a set of regional environmental regulatory authorities (known as CARs) charge all polluters a fee per unit of biological oxygen demand (BOD) and total suspended solids (TSS) discharged. According to Colombia's environment ministry, nationwide BOD discharges from point sources covered by the program fell 27 percent and TSS discharges fell by 45 percent (Blackman, Morgenstern & Topping 2006). However, these effects may not be causal. It is also important to note that these observed declines in water pollution may be due in part to more effective permitting, monitoring, and enforcement, as well as increased transparency and accountability of CARs (Blackman, Morgenstern & Topping 2006).

In the 1980s, Malaysia began charging a fee on BOD emissions from the palm oil industry. Unpublished studies suggest that the implementation of these fees reduced the BOD load dramatically, even as palm oil production increased (Vincent et al. 1997).

4.3.2 Tradable Pollution Permits

Tradable pollution permits, or "cap-and-trade" systems, are another market-based approach. Here the regulator sets an aggregate cap on pollution and allocates the number of pollution permits implied by the cap to the regulated community, either through auctions or a system of free allocation. The pollution permits are transferable, and thus when the permit market clears, each firm has equated its marginal pollution abatement cost with the prevailing permit price. This results in equal marginal costs across firms, which is the least-cost allocation of pollution control. When the marginal damages from water pollution vary with the location of the discharge, establishing location-based trading ratios (equivalent to a system of exchange rates) for each pair of polluters is an efficient approach (Konishi et al. 2015). Experience with water quality trading in industrialized countries is limited. Although there are active programs in Australia, Canada, and the United States, few are operating on a scale that could be considered economically significant (Fisher-Vanden and Olmstead 2013). To the best of our knowledge, none of these programs have been rigorously evaluated. We were not able to identify any examples of active water quality trading programs in developing countries. This is consistent with the fact that there are only a small number of tradable permit policies for air pollution in developing countries (Montero et al. 2002).

4.3.3 Payments for Ecosystem Services

Payments for ecosystem services (PES) are another type of market-based approach. When externalities create a divergence between those who bear the costs of pollution control and those who enjoy the benefits, a system of payments can potentially address these externalities. PES approaches aimed at controlling water pollution are known as payments for watershed services (PWS). As with taxes and tradable permits, efficient PWS systems must account for the spatial and intertemporal heterogeneity of marginal damages from pollution (Jack et al. 2008). A cost-effective PWS program maximizes the impacts of expenditures and avoids paying for abatement that would be undertaken even without a PWS incentive. This can be achieved through auctions (Ferraro 2008). According to Salzman et al. (2018), there are 550 active PES programs around the world.

4.3.3.1 Applications in industrialized countries

PWS approaches have been implemented for water pollution control in many industrialized country contexts. For example, in 2018, the US Department of Agriculture's Conservation Reserve Program (CRP) paid more than $1.8 billion to more than 300,000 US farms for environmentally-beneficial practices on 22 million acres (USDA 2018). Water quality improvement is an important CRP goal. Roberts & Lubowski (2007) find that the CRP results in the lasting retirement of agricultural land. Estimates of the causal impact of the CRP on water quality are not available, but empirical analysis by the

USDA suggests that the CRP generates tens of millions of dollars in recreational water quality benefits each year (Feather et al. 1999).

4.3.3.2 Applications in developing countries

The literature contains many reviews of PES applications in developing countries, including some robust PWS programs (Pagiola et al. 2005, Bulte et al. 2008, Pattanayak et al. 2010). We discuss some examples below. Note that we identified no PWS programs in developing countries that use auctions to select payment recipients, which suggests that there may be room to improve cost-effectiveness in these programs. In 2006, the city government of Beijing began paying farmers upstream of Miyun Reservoir to convert land from rice cultivation to dryland crops, in order to increase water yield in the catchment and reduce nutrient flows into the reservoir. Zheng et al. (2013) find evidence that this program, known as Paddy Land-to-Dryland, has been very successful, with an estimated benefit-cost ratio of 1.5 and net benefits flowing to both upstream service providers and downstream payees. In fact, by 2010, all rice fields upstream had converted to dryland crops (mostly corn), and concentrations of total nitrogen and total phosphorus (which are used in fertilizer) had been reduced significantly (Zheng et al. 2013).

Watershed protection and aquifer recharge are among the many goals of Mexico's federal conservation payments program. The program's main stated goal is to protect forests in order to maintain their "hydrological services". This program, which is financed by water user fees, has significantly reduced deforestation (American et al. 2015, Alix-Garcia et al. 2018), has increased land-cover management activities (Alix-Garcia et al. 2018), and does not appear to have crowded out private environmental stewardship (Alix-Garcia et al. 2018). While these induced changes may have improved water quality, we have not found any rigorous assessments of the program's water quality impacts.

Many smaller-scale PWS systems have been established in Latin America. For example, in Colombia's Chaina watershed in the eastern Andes, downstream water users pay upland farmers to switch to land-management practices that reduce soil compaction and erosion. Although no causal estimates of program impacts are available, Moreno-Sanchez et al. (2012) suggest that the program has both reduced deforestation and regenerated riparian vegetation, which could improve water quality. PWS programs in Bolivia's Upper Los Negros watershed and Ecuador's Palahurco watershed have also been discussed in the literature (Pattanayak et al. 2010). Kosoy et al. (2007) present case studies of three small-scale PWS programs in Honduras, Costa Rica, and Nicaragua, but the impacts of these programs on water quality have not been rigorously evaluated.

A PWS program has been piloted in Tanzania's Uluguru Mountains, which is the upland catchment area for the basin that provides water for most of Dar es Salaam and the surrounding regions. The Equitable Payment for Watershed Services program connects upland farmers with downstream water utilities, beverage companies (including Coca-Cola), and breweries (Mussa

& Mwakaje 2013). Water quality monitoring is occurring for this program, suggesting that it may be possible to measure its impacts (Branca et al. 2011).

4.3.4 Mandatory Information Disclosure

Mandatory information disclosure policies (e.g., requiring companies to publicly release information about their environmental performance) may correct a type of market failure – information asymmetry. The disclosure of information concerning a company's pollution emissions may affect consumers' demand for polluting firms' goods; firms' stock prices and their ability to hire and retain employees; private citizens' incentive to sue polluters; political support for more stringent pollution control standards or enforcement; and pressure from community groups and nongovernmental organizations. It may also provide new information to managers about plants' discharges and options for reducing them (Tietenberg 1998, Powers et al. 2011).

4.3.4.1 Applications in industrialized countries

Many information disclosure policies have been established and evaluated in industrialized countries. One of the most well-studied is the US Toxics Release Inventory (TRI) program, which requires manufacturing firms to report annual chemical releases into the air, water, and land to the US Environmental Protection Agency, which then publicly releases the information. Total annual releases of reportable chemicals fell by nearly 50 percent from the TRI's inception in 1986 through the mid-2000s. This decrease has not been causally attributed to the TRI (Bennear & Coglianese 2005), but studies of outcomes other than environmental performance have found that firms whose high TRI releases receive media coverage experience reduced stock returns (Hamilton 1995, Khanna et al. 1998). Similarly, the disclosure of environmental incidents and violations appears to have strong negative effects on the market value of firms in Canada (Laplante & Lanoie 1994) and Korea (Dasgupta et al. 2006).

4.3.4.2 Applications in developing countries

Information disclosure programs have often been used as environmental policy instruments in developing countries, although only a few have been rigorously evaluated. Indonesia's national Program for Pollution Control, Evaluation and Rating (PROPER) was created in 1995 to rate and disclose the environmental performance of factories (Tietenberg 1998). Early studies found that PROPER had a short-term impact on improving the performance of below-average firms but did not increase the number of firms using more than the required environmental management technologies (Tietenberg 1998, Blackman et al. 2004). More recently, García et al. (2007, 2009) evaluated PROPER's effectiveness and found that the program does reduce pollution emissions, especially for low-compliance firms.

India's Green Rating Project (GRP), which began in 1997, evaluates the environmental performance of large industrial plants in India, assigns numeric ratings to these plants, and awards them "leaves" to indicate their score. It also informs the public about the ratings and offers plants information about their pollution abatement options. In an evaluation of the impact of the GRP on discharges from India's largest pulp and paper plants, Powers et al. (2011) find that the program significantly reduced pollution from dirty plants.

4.4 Voluntary Approaches

Voluntary approaches (VAs) are alternative policy tools that do not fall into either the market-based or the prescriptive category. Under VAs, regulators either offer polluters incentives (e.g., cost-sharing programs, environmental leadership programs) to reduce pollution or induce participation by threatening stricter regulation if polluters do not adopt the VA (Borck & Coglianese 2009). The advantages of VAs over traditional regulations include: (1) potential cost savings, because polluters have the flexibility to choose abatement techniques to achieve environmental targets (as under a market-based approach); and (2) increased cooperation and communication between polluters and regulators (Alberini & Segerson 2002). However, a clear potential downside is that firms may be unlikely to engage in costly pollution reduction unless there are specific requirements, monitoring, and enforcement. In addition, when evaluating such programs, empirical analyses must address the issue of selection bias, because the firms most likely to join the VA are those for whom it is the least costly.

4.4.1 Applications in Industrialized Countries

VAs have been applied in several cases in the United States, Europe, and Japan, at both the federal and state levels. The evidence has been mixed on the effectiveness of one well-studied VA in the United States – the 33/50 program, which sought to achieve major reductions in releases reported under the TRI during the 1990s. For example, Vidovic & Khanna (2012) find no statistically significant decrease in pollution attributable to the program, while Khanna & Damon (1999) and Innes & Sam (2008) attribute significant reductions in releases to participation in the 33/50 program. These mixed results may be due to the fact that most empirical assessments of 33/50 do not account for the possibility that information about cost-effective abatement may spill over from enrolled to non-enrolled facilities. Zhou et al. (2020), who examine such spillovers directly, find that firms that do not participate in 33/50 may still reduce emissions, and that accounting for this possibility significantly increases the emissions reductions attributable to the program. Overall, it is not clear whether VAs actually improve water quality in industrialized countries or how they compare with mandatory regulations (Borck & Coglianese 2009).

4.4.2 Applications in Developing Countries

Several studies have examined VAs in developing countries, including those in Chile, Mexico, Colombia, China, and Brazil (Blackman & Sisto 2006, Blackman, Lyon & Sisto 2006, Jiménez 2007, Hu 2007, Blackman et al. 2010, 2013). However, only the study on Chile (Jiménez 2007) provides credible evidence that a VA improved environmental outcomes. Blackman et al. (2010) analyze the effectiveness of Mexico's Clean Industry Program, under which industrial facilities that agree to a third-party audit and implement the recommended changes can avoid penalties for any violations uncovered during the audit; they can also avoid government regulatory inspections for an additional two years. The results suggest that dirty firms recently punished by the government were more likely to participate in the program, but that after firms had met the requirements of the program, their pollution was not significantly lower than the pollution levels for nonparticipants. Blackman et al. (2013) find that VAs had minimal short-run effects on firms' environmental performance in Colombia. The authors argue that while empirical studies often fail to establish immediate impacts of VAs on environmental performance in developing countries, VAs may facilitate capacity-building in both government institutions and the private sector, which may help reduce pollution and improve regulatory outcomes over time.

4.5 Infrastructure Investment

Another tool that governments can use to reduce the impacts of water pollution is to directly fund or subsidize the construction of wastewater collection and treatment infrastructure or drinking water treatment and distribution infrastructure. An extensive literature suggests that the provision of safe drinking water and sanitation benefits households in many different ways. Although drinking water interventions can reduce economic damages from ambient water pollution (by weakening the link between ambient pollution and human health), we would not expect drinking water interventions to have any direct effect on ambient pollution itself. In contrast, major investments in sanitation infrastructure can improve ambient water quality and thus affect human and ecosystem health. Untreated waste is a classic negative externality, while centralized wastewater collection and treatment provide public goods. Although most positive externalities of piped water are likely to be internalized within a region (i.e., most of the benefits of piped water provision tend to accrue to individuals living within the jurisdiction that pays for it), wastewater treatment generates spillover benefits to downstream regions (Chiang 2016). Thus, there is a strong economic rationale for government provision of sanitation infrastructure, and large-scale investments are indeed common in both industrialized and developing countries.

In the case of industrialized countries, the literature has emphasized the role of water treatment and sewerage systems in the decline in mortality rates in US and European cities in the early 20th century, which had very substantial net economic benefits (Alsan & Goldin 2019, Cutler & Miller 2005, Delaney et al. 2011). The literature also finds positive impacts of toilets, latrines, and safe drinking water on health and educational outcomes in developing countries (Soares 2007, Duflo et al. 2015, Zhang 2012, Zhang & Xu 2016).

In developing country settings where the share of untreated sewage is large, and especially in urban areas with large exposed populations, large-scale sanitation investments could have significant impacts on downstream water quality as well as significant economic benefits via avoided morbidity and premature mortality (Kresch et al. 2020). However, the literature on the impacts of large-scale sanitation investments on downstream ambient water quality is generally very thin and mixed. As noted earlier, Greenstone and Hanna (2014) find that India's NRCP, which focuses on investments in wastewater collection and treatment (as well as community toilets, crematoria, and public education), has had no significant impact on ambient water quality. The dominant role of community sanitation improvements, rather than in-home improvements, in producing health benefits (e.g., Andres et al. (2017) suggests that some of the positive effects of sanitation investments could be due to ambient water quality improvements, but to our knowledge, this link has not been made directly. Thus, this is an important area for further research.

4.6 Challenges for the Design, Implementation, and Evaluation of Water Pollution Control Policies in Developing Countries

In this section, we highlight challenges to the effective design, implementation, and evaluation of water pollution control policies in developing country settings. These challenges include sparse and/or poor quality data, weak monitoring and enforcement, rent-seeking in regulatory settings, and jurisdictional spillovers when regulation is decentralized.

4.6.1 Data Availability

One major challenge for evaluating water pollution control policies in developing countries is a lack of data on outcome variables such as pollution emissions or pollution concentrations in receiving waters. One tool for addressing this challenge is the increasing availability of satellite data that measures pollution (Schaeffer et al. 2012). Another option for evaluating policy effectiveness in data-poor settings is for researchers to use variation in water pollution that is due to exogenous shocks that have direct connections to ambient pollution concentrations (e.g., policy interventions, natural events), but that do not require direct pollution measurement. For example, Do et al. (2018) exploit variation in industrial pollution on India's Ganga River (due to Supreme Court rulings that mandated pollution reductions from the tanning industry) to estimate the impacts of pollution reductions on neonatal mortality; they do not use any direct measures of river water quality. Field experiments such as randomized control trials can also be used to evaluate water and sanitation interventions without needing to rely on observational data provided by regulators or other sources (Kremer et al. 2011). Such experiments have been successfully used to examine the impacts of specific regulatory interventions on the emissions of individual firms or households, which certainly contribute to ambient water pollution concentrations (e.g., Duflo et al. (2013)). But it would be problematic to extrapolate from these relatively small-scale causal estimates to large-scale ambient water pollution reductions. Innovative use of available monitoring technologies and increased monitoring frequency could improve data availability and quality in developing countries. One recent example of such an effort is China's automatic water quality monitoring system. Data quality has been a significant concern for studies of pollution in China (Ghanem & Zhang 2014). China's Ministry of Ecology and Environment (MEE) (formerly the Bureau of Environmental Protection) has collected data on major rivers and lakes using an automatic water quality monitoring system since 1999, with monitoring stations collecting and publishing real-time water quality data online. Compared to previous monitoring regimes, it may be more difficult to manipulate these data. China's central government recently expanded the automatic water quality monitoring system from 100 to 2,050 stations across the country (Xinhua Net 2018).

Third-party monitoring can also improve water quality data in developing countries. For example, Duflo et al. (2013) find that in India, having a paid third-party auditor who is externally selected results in more accurate reporting of firms' water pollution emissions. Citizen science, whereby individuals voluntarily participate in the data collection and monitoring process, can also function as a form of third-party monitoring. However, thus far, the literature provides only theoretical support for the idea of such crowdsourcing of water quality data (Borden et al. 2016).

4.6.2 Inadequate Monitoring, Enforcement, and Compliance

There is a large environmental economics literature that examines the monitoring and enforcement of pollution control regulations (Shimshack 2014).

Indeed, dozens of studies have found that monitoring and enforcement of environmental regulations reduce pollution, deter future violations, and even encourage over-compliance by regulated entities (Shimshack 2014). Most of these studies have focused on the United States, likely because of both data availability and the country's long history of pollution regulation. Many studies have examined the weak institutional capacity for stringent monitoring and enforcement in developing countries and the challenges this poses (Afsah & Makarim 1999, Wang & Wheeler 2000, Dasgupta et al. 2000). Although it is not unusual for developing countries to have environmental standards that are actually quite stringent, they are often not met simply because enforcement is weak (Greenstone and Jack 2015). A few empirical studies have focused specifically on the importance of monitoring and enforcement of water pollution policies in developing countries. For example, Dasgupta et al. (2001) show that in China, inspections may play a more significant role in reducing emissions than the pollution levies themselves. **?** find that the impacts of China's pollution levies are higher in areas of the country where regulatory institutions (and hence monitoring and enforcement) are stronger. Using plant-level pollution data, Lin (2013) shows that inspections increase plants' self-reported pollution by more than 3 percent but may not result in reduced pollution. In another study of China, Zhang et al. (2018) assess the central government's National Specially Monitored Firms (NSMF) pilot program, which was established in 2007 and oversees the local monitoring of firms that are major emitters of air and water pollution and hazardous waste. More specifically, monitored firms are required to install automatic monitoring systems and transmit emissions information in real time to the central government, which verifies the accuracy of the data through monthly inspections. Zhang et al. (2018) find that the additional central supervision of local authorities in the NSMF program reduced industrial COD emissions by 26.8 percent in the first year of the program, with reductions continuing in subsequent years.

4.6.3 Rent-seeking and Environmental Regulation

Rent-seeking under regulatory systems can drive a wedge between what policy makers expect an environmental policy to achieve and actual results (Wilson & Damania 2005). For example, firms may simply lobby for less stringent environmental regulation or regulators may be "captured" (be subject to influence) by the industries they monitor. A small number of studies explicitly examine the effects of corruption and related rent-seeking behavior on water pollution outcomes in developing countries. For example, Duflo et al. (2013) find that in India, aligning environmental auditors' and regulators' incentives (by randomly assigning auditors to plants, paying auditors from a central pool rather than by plant, and verifying their reports with follow-up inspections) significantly reduces under-reporting of water pollution emissions from audited plants, although the effects on reporting are somewhat smaller for water pollution than for air pollution. The firms in this experiment actually reduced water pollution emissions in response to better auditing, but air emissions showed no statistically significant impact (Duflo et al. 2013).

4.6.4 Decentralized Regulation and Inter-jurisdictional Spillovers

When regulation is decentralized, local jurisdictions are able to set standards that reflect local preferences for environmental quality, and individuals may sort across jurisdictions according to their preferences for environmental quality and other factors (Tiebout 1956, Oates & Schwab 1988). However, because water pollution moves across political and geographic boundaries, it creates negative externalities (in which pollution spills over from the originating jurisdiction to jurisdictions downstream) and free-riding problems. Polluting facilities may even move in response to regulatory stringency within a jurisdiction, offsetting any water quality improvement within the more stringent jurisdiction with increases in pollution elsewhere (Decheleprêtre and Sato 2017). Although centralized regulation is less responsive to local preferences for environmental quality, it can internalize spillovers by considering all impacts of pollution and pollution control, regardless of their location. Which regulatory approach – centralized or decentralized – is most efficient depends on whether the efficiency loss from centralized standard-setting (which ignores local conditions) exceeds the efficiency gain from eliminating the spillovers that occur in decentralized policy settings (Banzhaf & Chupp 2012). In the remainder of this discussion, we examine the empirical evidence concerning the existence of free-riding in water pollution across jurisdictional borders, as well as potential policy solutions.

4.6.5 Empirical evidence of spillovers

There is some empirical evidence of water pollution spillovers both within and across country borders in industrialized country settings (e.g., Sigman (2005)). Empirical studies suggest that such transboundary spillover and free-riding problems also exist in developing countries. For example, in 2001, the Chinese government mandated pollution reduction targets for all provinces in its Five-Year Plan, but did not indicate how local Bureaus of Environmental Protection (BEPs) should coordinate to achieve these targets. Because local BEPs are controlled by local governments, they had an incentive to strategically place polluting industries in border counties. Cai et al. (2016) assess the impacts of this approach and find that water-polluting production and new entry into water-polluting industries are significantly higher downstream of county borders, providing evidence of the jurisdictional spillovers described above. Thus, in response to the 2001 policy directive, provincial governments appear to have allocated the most lenient enforcement to downstream border counties (Cai et al. 2016). Similarly, Chiang (2016) finds that local Chinese government officials who receive incentives from the central government to expand the fraction of households within their jurisdictions that are covered by clean piped water and sanitation systems do more to expand access to piped water than to expand access to sanitation. This is because sanitation expansion creates positive spillovers to downstream jurisdictions (via reduced ambient water pollution) that are not captured locally. In a study of Brazil, Lipscomb & Mobarak (2017) examine the potential water pollution spillovers in rivers as they approach county borders. They find that pollution increases as rivers travel towards downstream borders and that the rate of pollution increases as rivers approach a border. As with the studies of China, these findings provide evidence of the jurisdictional pollution spillovers that result from decentralization.

4.6.6 Addressing the problem of spillovers

While there is now clear evidence of the impact of inter-jurisdictional spillovers on water quality, there are no clear solutions to the problem. Theory suggests that under certain restrictive circumstances, private negotiation among actors can provide efficient solutions to such negative externalities (Coase 1960, Anderson & Libecap 2014). Coasian solutions appear to have some potential in developing countries. For example Kremer et al. (2011) show that privatization of communal property rights to local springs in Kenya (i.e., internalizing the benefits of spring protection) could increase welfare as incomes rise, costs of protection fall, or water becomes more scarce. However, such private solutions to ambient water pollution are unlikely to arise in an international context where there is no binding legal framework to facilitate negotiation and enforce contracts. Another possible solution to the problem of inter-jurisdictional spillovers is to provide stronger incentives to local regulators to improve water quality in contexts where regulation is decentralized. Three studies have examined the impacts of such an approach in China, where, beginning in 2005, the central government included water quality as a criterion for promotion of local government officials. Kahn et al. (2015) find that while COD decreased significantly at local boundaries after this policy change, emissions of more harmful pollutants, such as petroleum and mercury, were not affected. In a follow-up study, Chen et al. (2018) show that the policy has had unintended consequences. After the policy change, water quality in the Yangtze River deteriorated despite all provinces having achieved their COD reduction goals. This is because upstream provinces of the Yangtze River are less economically developed and have relatively better water quality, and thus also have less stringent COD reduction targets; as a result, water-polluting industries have shifted to less-regulated areas upstream (Chen et al. 2018). Local government officials also appear to have enforced tighter standards on polluting firms immediately upstream of monitoring stations after the central government explicitly linked political promotion to water quality improvements (He et al. 2020).

4.7 Conclusions and Research Gaps

Our review of the literature on water pollution control policies in developing countries provides a map of the gaps in the evidence on policy effectiveness, thus highlighting important topics for future research. Overall, we identified only about a dozen plausibly causal estimates of the impacts of specific pollution control policies on either water pollution emissions or ambient water pollutant concentrations in developing countries. The majority of these studies focus on water pollution policies in China. The volume of literature assessing the impacts of policies on water pollution emissions or ambient concentrations is similarly thin for developing countries and industrialized countries. The literature evaluating other water pollution policy impacts, such as those on human health and educational outcomes, is even thinner. Given the low number of studies with plausibly causal estimates and the fact that the evidence for human health and other impacts is quite convincing in those studies that have been published (e.g., Ebenstein 2012, Do et al. 2018, Garg et al. 2018), our first key finding is that there is a a great need for rigorous evaluation of water pollution control policies in the literature.

A second key finding is that the theory of market-based approaches to environmental policy is well-established and applications of taxes, tradable permits, and information disclosure policies in industrialized and developing countries have demonstrated that market-based approaches can achieve pollution reduction goals. However, only two studies in the literature rigorously evaluate the cost-effectiveness of these approaches in a developing-country setting (Ebenstein 2012, Zheng et al. 2013). Given the scale of water pollution problems in developing country rivers, lakes, streams and coastal waters, market-based pollution control policies are a promising tool for countries to consider. Indeed, recent benefit-cost analyses of primarily prescriptive approaches in the United States suggest that their costs may exceed their benefits (Keiser & Shapiro 2019*c*), highlighting the need to focus on cost-effective approaches.

Unfortunately, developing countries have few concrete examples on which to draw when designing new market-based water pollution control policies. China's pollution levy system provides one good example, with multiple analyses suggesting that it has reduced water pollution emissions. The newer pilot tax policy in China's Lake Tai Basin is an even better example, given that the tax is assessed on all units of pollution, and it appears to have dramatically reduced water pollution. Perhaps it is not surprising that we have uncovered so few clear examples of successful market-based policies for water pollution control in developing countries; the record in industrialized countries is also thin.

A third important finding of our review is that PWS and information disclosure policies have been applied to water pollution control in developing countries more often than one might expect, perhaps even exceeding their use in industrialized countries. However, there are too few rigorous evaluations of their impacts to draw clear conclusions about the effectiveness of these policies for improving ambient water quality.

Fourth, the literature suggests that major public infrastructure investments, particularly in municipal wastewater treatment, clearly paid off for industrialized countries during the early stages of their development. Given the evidence, it seems likely that major urban wastewater treatment investments would be net beneficial in developing countries, but projects and investments would need to be evaluated individually.

Finally, we note that where regulatory institutions, monitoring, and enforcement are weak, the impacts of water pollution control policies are also likely to be weak. The challenges of decentralization and water pollution spillovers must also be considered in developing country settings. On a positive note, recent experiments with industrial pollution auditors in India and econometric analysis of policies that increase central oversight of local pollution monitors in China provide evidence that fairly simple, low-cost interventions may be effective in overcoming some of the challenges that are specific to water pollution policy in developing countries.

Chapter 5

Conclusion

5.1 Conclusion and contribution

This dissertation contributes to the literature by examines three important questions on the economics of water pollution control.

In the first paper, I investigate the benefit of water pollution abatement using an integrated two-stage model of hedonic analysis and RUM model. I find that increases in dissolved oxygen (DO) improve both recreational and aesthetic amenities and that homeowners in Tampa Bay have significant MWTP for both of these improvements. Using data from the Tampa Bay Estuary Program on the cost of nutrient pollution reduction projects from 1998-2014, I find very favorable benefit-cost ratios when compared to other water quality benefit-cost analyses in the literature (Keiser & Shapiro 2019*b*,*a*, Keiser et al. 2019).

This work adds to our understanding of how people value water quality improvements, especially nutrient pollution abatement. Eutrophication, a consequence of nutrient pollution, may cause large economic damages in the United States and elsewhere. Many local, state, and federal regulations have been implemented to address this problem. Further work to help policymakers better understand how people value nutrient pollution abatement, and how these values are capitalized in housing markets, can contribute to a more comprehensive evaluation of such regulations.

This paper also contributes to the literature on hedonic valuation of pollution control, more generally. We estimate the first hedonic model valuing water quality that controls comprehensively and flexibly for property characteristics, using two different approaches. Our long-difference hedonics approach may comport better with hedonic theory than other approaches in the literature, given that the hedonic model considers the property location decision in long-run equilibrium. Lacking data on recreation site visitation at the property level—likely a problem faced by other researchers examining similar questions, unless they implement a household survey—we use multiple imputations to address the resulting measurement error relative to recreation data observed by property. These innovations may enable future work valuing water pollution and pollution control with a broader geographic scope than we have examined in this paper.

In the second paper, using restrictive education data and drinking water data in Texas from 1991 to 2011, I investigate the cost of an important drinking water pollution in the United States by estimating the educational impacts of drinking water lead exposure in early childhood. My results provide empirical evidence that drinking water lead exposure in the birth year has significant negative impacts on standardized test scores in the third grade. The magnitudes of my coefficients are similar to that estimated for blood lead levels in the existing literature (Aizer & Currie 2019). Importantly, the effects I estimate occurred in a state and at a time when the lead concentration in drinking water is, on average, quite low, and the magnitude (on average) below the federal action threshold under the SDWA. I also find that violating the federal regulation of drinking water lead treatment in the birth year leads to a smaller probability of graduating from high school. Given the income premium of having a high school diploma, an LCR violation at birth may be associated with a foregone wage of $12 million annually in Texas alone. The associated life earning the loss in lifetime earnings range around $241 million to $1.6 billion. Lastly, I find that female students and children from African American families and families with economic disadvantages are more vulnerable to lead exposure in line with existing literature (Chetty et al. 2016, Heckman & Karapakula 2019, Grönqvist et al. 2020). The findings suggest that early childhood lead exposure may be on contributing factor to the gender-achievement gap and racial-achievement gaps in the U.S.

This paper contributes to the literature by providing the first evidence of the impacts of contemporary U.S. drinking water lead levels on elementary school test scores. It also provides the first evidence that these effects persist through longer educational milestones, such as high school graduation, in the United States. Third, this paper contributes to the literature on lead exposure's implications for inequality. Lastly, it uses the surface water Cl^- concentration as an instrument for drinking water lead concentration. This instrument was not used before in the economics literature.

In the third paper, I reviewed the literature on water pollution control policies in developing countries and provided a map of the gaps in the evidence on policy effectiveness. Overall, I find there is a great need for rigorous evaluation of water pollution control policies in the literature given the lack of empirical evidence on the effectiveness of water quality control policies. Second, the evidence on the cost-effectiveness of market-based approaches in both developed and developing country contexts needs more research. A third important finding of the review is that PWS and information disclosure policies have been applied to water pollution control in developing countries more often but there are too few rigorous evaluations of their impacts to draw clear conclusions about the effectiveness of these policies for improving ambient water quality. Fourth, the literature suggests that major public infrastructure investments, particularly in municipal wastewater treatment, but the evidence on the effectiveness of such investments on water quality in developing countries need to be studied on a project basis. Finally, the literature review notes that where regulatory institutions, monitoring, and enforcement are weak, the impacts of water pollution control policies are also likely to be weak. The challenges of decentralization and water pollution spillovers must also be considered in developing country settings.

5.2 Future Research

This dissertation studied three important aspects of the cost-benefit analysis on water pollution control. My future research will include:

The first study using data in Texas provides the first empirical evidence that early childhood exposure to lead from drinking water has a significant impact on educational outcomes in both the short and the long run. However, my IV sample of 5 Texas counties and the comparably lower lead concentration in Texas may not provide a comprehensive estimate of drinking water lead impact on the whole U.S. An important direction of my future research is to take the approach and apply it to a broader geographic area. It serves as a pilot for a larger project to examine the impact of lead in drinking water across the whole United States. I have recently gained access to the Federal Census Restrictive Data Center to use the restrictive census data for my project for up to 5 years. The federal restrictive data will enable future work to evaluate drinking water lead impacts with a broader geographic scope than I have examined in this paper.

In the second paper, collaborated with Dr. Olmstead and Dr. Kuwayama, though our study only focuses on one city in the United States, our estimation suggests that the omission of recreational benefits may be the reason for the lack of evidence that the benefits of water quality improvement exceed the cost of water pollution control. We are also interested in follow-up with this project and apply the approach to the whole United States, with the goal of providing a more comprehensive estimate of the value of U.S. water quality improvements under the Clean Water Act.

I would like to build on the existing literature review I have done and examine the costs and benefits of water pollution control in the context of developing countries. Developing countries' context poses unique challenges for estimating the benefits of water quality improvements. For instance, many developing countries may lack a well-functioning housing market, which is a basic assumption of the hedonic model. Since my co-author and I examine the empirical evidence for the effectiveness of water pollution control policies in developing countries and find that evidence regarding market-based policies is quite thin in the third paper, I would like to examine policy evaluation of water pollution control in developing countries and its impact on health and labor market participation. One example of my future research direction is my sole-authored working paper on China's River Chiefs program. This paper examines how the River-Chief policy in China affects surface water quality, seeking to understand whether assigning local officials responsibility for river segments in their jurisdiction contributes to reducing water pollution and the negative externalities of pollution to a jurisdiction's downstream neighbors. As surface water pollution may reduce agricultural productivity in developing countries like China, one negative externality of surface water pollution I will test for in this paper is the health and educational outcomes of individuals in downstream neighboring cities.

Appendices

Appendix A

Chapter 3 Appendix

A.1 Appendix for Online Publication: Additional Figures and Tables

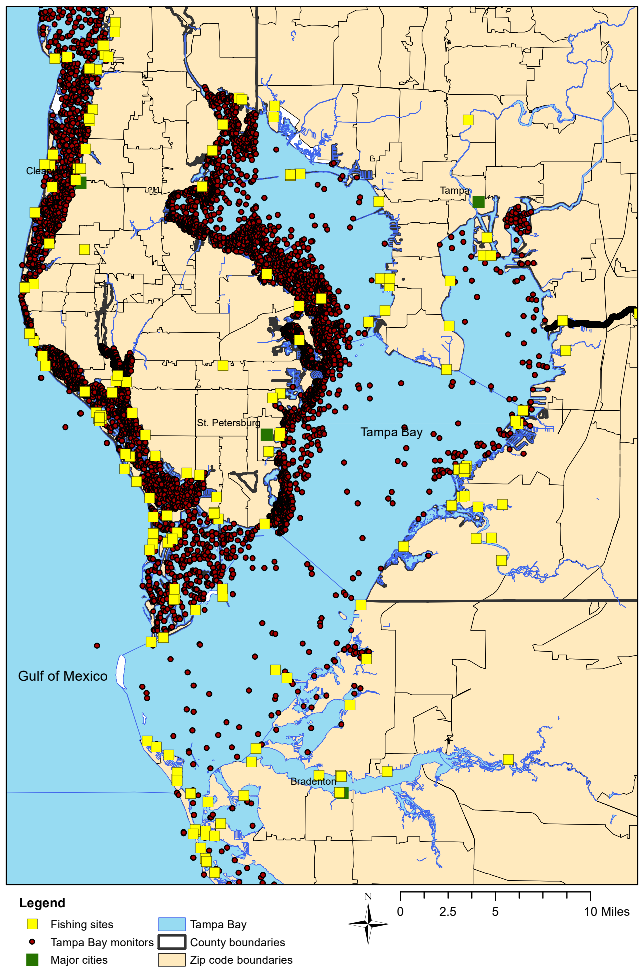


Figure A.1: Location of fishing sites and Tampa Bay water quality monitors

Notes: Water shapefiles data from the Tampa Bay Water Atlas website. We use the Gulf of Mexico and Bay waterbodies to define Tampa Bay. We then spatially join water quality monitors from STORET with the Tampa Bay shapefile to define Tampa Bay monitors. Fishing site locations are from the MRIP dataset, and zip code boundaries are from the U.S. Census Bureau.

Table A.1: Additional water quality descriptive statistics

	Dissolved oxygen
mean	5.94026
min	0
5th percentile	1.68
95th	8.89
max	28740
# of	
obs (without missing)	209336
monitoring sites	5913
mean readings per monitor per year	53
mean readings per monitoring site	443
mean years per monitoring site	8
missing	44
yearly average	
# of obs	22714
mean	5.92772
min	0
5th percentile	2.5265
95th percentile	8.725
max	10.8

Table A.2: Summary statistics by DO level in nearby water

Variable	DO $\geq$ 5mg/L	DO<5mg/L	Full sample
DO level	6.786	3.995***	5.792
	(4.76)	(0.781)	(4.075)
ECS_{jt}	35.88	36.16***	35.98
	(2.64)	(1.96)	(2.42)
Property age	32.16	34.48***	32.99
	(20.44)	(22.08)	(21.07)
Price (2014 dollars)	235534.5	218013.8***	229306.5
	(158425.1)	(147162.2)	(154742.5)
Distance to local water	896.36	987.95***	928.92
	(1052.92)	(1004.02)	(1036.73)
Distance to Tampa Bay	14479.78	16836.08***	15317.35
	(13633.68)	(17619.91)	(15212.89)
Local water front	0.046	0.046	0.046
	(0.209)	(0.209)	(0.209)
Tampa Bay front	0.0107	0.011	0.0109
	(0.103)	(0.105)	(0.104)
N	94,684	52,219	146,903
Hillsborough County			
Number of bedrooms	3.208	3.131***	3.17
	(0.806)	(0.815)	(0.811)
Number of bathrooms	2.095	2.042***	2.069
	(0.671)	(0.703)	(0.687)
Number of stories	1.181	1.172***	1.177
	(0.408)	(0.407)	(0.407)
Heated area	1764.56	1697.05***	1731.83
	(674.65)	(663.30)	(670.02)
Lot acreage	0.292	0.230***	0.267
	(0.478)	(0.269)	(0.407)
N	33,641	31,660	65,301

Notes: Means, with standard deviations in parentheses, for observations used in regression analysis. Asterisks in column 2 indicate significant difference in means between the two groups, according to a t-test for difference in means.
* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.3: First-stage recreation demand model with DO dummy

	(1) Travel cost (3km)	(2) Travel time (3km)	(3) Travel cost (5km)
Travel cost (US dollars)	-0.110***		-0.113***
	(0.00078)		(0.00080)
Travel time (minutes)		-0.0639***	
		(0.00044)	
DO > 5mg/L	0.184***	0.208***	-0.034
	(0.0289)	(0.0287)	(0.0376)
Seagrass abundance	-0.164***	-0.128***	-0.139***
	(0.0104)	(0.0103)	(0.0104)
Alternative-specific constants	Yes	Yes	Yes
Observations	1,765,796	1,765,796	1,801,615

Standard errors in parentheses.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Notes: Models are estimated using conditional logit, with a choice set of 85 fishing sites visited during the study period. Columns 1 and 2 link fishing sites to Tampa Bay water quality monitors within 3km. Column 3 links sites with monitors within 5km. Travel cost is estimated as the sum of the value of travel time (1/3 of foregone wages times round-trip travel time) and the operational cost of travel (AAA's driving cost times round-trip distance).

Table A.4: Hedonic results without property fixed effects (Hillsborough County)

	(1) Basic 3km	(2) No ECS_{jt} 3km
ln(DO)	-0.00293	
	(0.00295)	
DO $\geq$ 5mg/L		-0.00330
		(0.00298)
ECS_{jt}	-0.314***	-0.310**
	(0.0889)	(0.101)
Property age	-0.00338***	-0.00339***
	(0.000172)	(0.000174)
Lot acreage	0.0282***	0.0288***
	(0.00747)	(0.00762)
Heated area	0.000563***	0.000563****
	(0.00000669)	(0.00000626)
Number of bedrooms	-0.0296***	-0.0295***
	(0.00425)	(0.00424)
Number of bathrooms	0.102***	0.101***
	(0.00568)	(0.00574)
Number of stories	-0.0643***	-0.0642***
	(0.00686)	(0.00680)
Property FE	No	No
Year FE	Yes	Yes
N	65,301	65,301
R-squared	0.704	0.704

* p<0.10, ** p<0.05, *** p<0.01

Notes: Dependent variable is the log property transaction price. Both columns use a 3-km radius to define average water quality around properties. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1. Standard errors are clustered by property.

Table A.5: Second-stage hedonic regression results with DO dummy

	(1) Basic 3km	(2) No ECS_{jt} 3km	(3) Basic 5km	(4) Travel time for ECS_{jt}	(5) County time trend	(6) Subdiv. time trend
DO $\geq$ 5 mg/L	0.00729***	0.00722***	0.00468*	0.00746***	0.00452*	0.00243
	(0.00256)	(0.00257)	(0.00252)	(0.00256)	(0.00256)	(0.00272)
ECS_{jt}	0.264***		0.161**	0.115**	0.0639	0.0267
	(0.0827)		(0.0810)	(0.0453)	(0.0773)	(0.0866)
Property age	-0.0122***	-0.0122***	-0.0104***	-0.0122***	-0.0135***	-0.0139***
	(0.00334)	(0.00334)	(0.00314)	(0.00334)	(0.00332)	(0.00342)
Property FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County year trend	No	No	No	No	Yes	No
Subdivision year trend	No	No	No	No	No	Yes
N	146,903	146,903	183,582	146,903	146,903	125,276
R-squared	0.626	0.626	0.627	0.626	0.632	0.631
MWTP for local DO $\geq$ 5 mg/L ($)	1,672	1,656	1,073	1,711	1,036	557
MWTP for Tampa Bay DO $\geq$ 5 mg/L ($)	101,227	N/A	61,734	44,096	24,502	10,123

Standard errors in parentheses.

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: The dependent variable is the log property transaction price. Column 1 uses a 3-km radius to define average water quality around properties. Column 2 drops the recreational utility index, ECS_{jt}. Column 3 repeats column 1, using a 5-km instead of a 3-km radius to define average water quality around properties. N rises in column 3 because more repeat sales are located within 5 km of at least one water quality monitor than within 3 km. Column 4 uses travel time instead of travel cost in the first stage to estimate ECS_{jt}. Column 5 includes county*year trends as additional controls. Column 6 includes census subdivision*year trends. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in detail in Section 5.2.1.

Table A.6: Second-stage long difference models with varying definitions of period a and period b

	(1) Longer time SE property	(2) Longer time SE zipcode	(3) Half time SE property	(4) Half time SE zipcode
$\Delta ln(DO)$	0.0295***	0.0295	0.0273***	0.0273
	(0.00764)	(0.0267)	(0.00767)	(0.0264)
ΔECS	0.0241***	0.0241***	0.0240***	0.0240***
	(0.00142)	(0.00281)	(0.00141)	(0.00285)
ΔProperty age	0.0585***	0.0585***	0.0573***	0.0573***
	(0.00190)	(0.00425)	(0.00186)	(0.00418)
Property FE	Yes	Yes	Yes	Yes
Year FE	No	No	No	No
Cluster SE level	Property	Zipcode	Property	Zipcode
N	27,452	27,452	34,571	34,571
R-square	0.272	0.272	0.177	0.177
MWTP for 1 mg/L local DO ($)	980	980	907	907
MWTP for 1 mg/L Tampa Bay DO ($)	3,083	3,083	3,070	3,070

* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Notes: Dependent variable is the log property price. Columns 1 and 2 define period a as 1998-2006, and period b as 2009-2014. Columns 3 and 4 define period a as 1998-2007, and period b as 2008-2014. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1. Robust standard errors are clustered by property or by zip code, as labeled in each column.

Table A.7: Hedonic model with proximity to water

	(1) Property FE	(2) Long difference
ln(DO)	0.0160***	
	(0.00515)	
ln(DO) $\times$ Distance to local water	-0.00217	
	(0.00160)	
ECS_{jt}	0.354***	
	(0.0832)	
ECS_{jt} $\times$ Distance to Tampa Bay	-2.06e07***	
	(2.74e-08)	
Property age	-0.0123***	
	(0.00331)	
$\Delta ln(DO)$		0.0309***
		(0.00536)
$\Delta ln(DO)$ $\times$ Distance to local water		-0.00319
		(0.00366)
ΔECS_j		0.0119***
		(0.00183)
ΔECS_j $\times$ Distance to Tampa Bay		1.98e-09
		(6.13e-08)
Δ Property age		0.0351***
		(0.00178)
Property FE	Yes	No
Year FE	Yes	No
N	146,903	14,390
R-Squared	0.627	0.049

Robust standard errors in parentheses are clustered by property.
* p<0.10, ** p<0.05, *** p<0.01.
Notes: Column 1 includes interactions of ln(DO) with the distance from a property to local water and of ECS_{jt} with the distance from a property to Tampa Bay. Column 2 is the long difference model with the same interaction terms. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1.

Table A.8: Hedonic model and long difference model with park water quality monitors

	(1) Hedonic model	(2) Long difference model
ln(DO)	0.0185	
	(0.0123)	
ECS_{jt}	1.161***	
	(0.221)	
ln(parkDO)	-0.0252**	
	(0.0117)	
Property age	-0.109***	
	(0.0223)	
$\Delta ln(DO)$		0.000598
		(0.0183)
ΔECS_j		0.0209***
		(0.00228)
$\Delta ln(parkDO)$		0.177***
		(0.0222)
Δ Property age		0.0410***
		(0.00313)
Property FE	Yes	No
Year FE	Yes	No
N	41,618	3,590
R-squared	0.622	0.054

* p<0.10, ** p<0.05, *** p<0.01.

Notes: Dependent variable is the log property transaction price. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1. Robust standard errors are clustered by property.

Table A.9: Estimated coefficients for local DO using smaller radii for monitors

	(1) Property FE Continuous DO	(2) Property FE DO ≥ 5 mg/L		(3) Long difference Continuous DO
5km monitors	-7.03e-06	0.00422*	5km monitors	0.0486***
(N=183,582)	(0.00390)	(0.00251)	(N=19,210)	(0.00999)
3km monitors	0.0114***	0.00729***	3km monitors	0.0226***
(N=146,903)	(0.00307)	(0.00256)	(N=14,390)	(0.00645)
1km monitors	0.0158*	0.0187***	1km monitors	0.00274
(N=32,996)	(0.00865)	(0.00587)	(N=2,515)	(0.0316)
500m monitors	0.0249***	0.0282***	500m monitors	-0.0430
(N=18,403)	(0.00956)	(0.00743)	(N=1,438)	(0.0342)
300m monitors	0.00623	0.0235	300m monitors	0.0222
(N=3,056)	(0.0216)	(0.0178)	(N=205)	(0.0707)

* p<0.10, ** p<0.05, *** p<0.01.

Notes: Dependent variable is the log property transaction price in columns 1 and 2, and the log long-difference in price in column 3. Only lnDO coefficients are reported, but columns 1 and 2 contain the same covariates as Table 2.3, column 1, and column 3 contains the same covariates as Table 2.4, column 1. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1. Robust standard errors are clustered by property.

Table A.10: Hedonic regression results using DO moving averages

	(1) 1 year	(2) 6 months	(3) 3 months
ln(DO)	0.0106***	-0.000246	-0.00796**
	(0.00350)	(0.00347)	(0.00319)
ECS_{jt}	0.154**	0.163**	0.174**
	(0.0677)	(0.0743)	(0.0833)
Property age	-0.0121***	-0.0122***	-0.0122***
	(0.00329)	(0.00329)	(0.00329)
N	162,765	147,489	133,027
R-square	0.604	0.610	0.612

* p<0.10, ** p<0.05, *** p<0.01.

Notes: Dependent variable is the log property transaction price. Columns 1-3 average local and recreational water quality 12 months, 6 months, and 3 months prior to each property transaction, respectively. Reported coefficient estimates and standard errors are obtained using multiple imputation, using methods described in Section 5.2.1. Robust standard errors are clustered by property.

Appendix B

Chapter 2 Appendix

Table B.1: Second IV estimates of lead impact using LIML

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A	Reading scores					Math scores				
Lead * pipes	-0.910***	-2.192***	-2.304***	-1.453	-3.911**	-2.416***	-3.537***	-4.093***	-5.637***	-5.198***
	(0.061)	(0.039)	(0.121)	(0.886)	(1.956)	(0.095)	(0.040)	(0.063)	(0.520)	(2.345)
B	Meet reading standard					Meet math standard				
Lead * pipes	-0.001***	-0.005***	-0.006***	-0.006***	-0.0137***	-0.007***	-0.010***	-0.011***	-0.0174***	-0.0127***
	(0.0003)	(0.00008)	(0.0005)	(0.0009)	(0.005)	(0.0002)	(0.00005)	(0.0002)	(0.001)	(0.00287)
Observations	361106	361024	361106	361106	361106	361106	361024	361106	361106	361106
First stage F statistics	3748.53	3860.22	1761.13	529.17	30.86	3748.53	3860.22	1761.13	529.17	30.86
Birth county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County by year trend	No	No	No	No	Yes	No	No	No	No	Yes
Individual controls	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Neighborhood controls	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Weather controls	No	No	No	Yes	Yes	No	No	No	Yes	Yes

Standard errors in parentheses and clustered at county level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix C

Chapter 4 Appendix

C.1 Appendix Tables

Table C.1: Empirical evidence on the impacts of water pollution control policies in developed countries

		Environmental impacts			Health impacts	
Water pollution control policies		***Reduction in emissions/effluent***	***Ambient water quality improvement***	***Land use change***	***Reduced morbidity***	***Reduced premature mortality***
Prescriptive policies	***Technology standards***		Keiser and Shapiro 2018			
	Performance standards	Magat & Viscusi 1990; Laplante & Rilstone 1996; Shimshack & Ward 2005, 2008; Earnhart 2004a, 2004b; Glicksman & Earnhart 2007; Shimshack 2014	Keiser and Shapiro 2018			
Market-based policies	***Pollution taxes***	Bressers 1998; Glachant 2002; Boyd 2003				
	Tradable permits	Fisher-Vanden and Olmstead 2013; Environment Protection Authority, State of New South Wales 2018; Connecticut Department of Energy & Environmental Protection 2016	Fisher-Vanden and Olmstead 2013			
	Information disclosure	Khanna et al. 1998; Bae et al. 2010				
	Payments for ecosystem services	Appleton 2002; Depres et al. 2008; Salzman et al. 2018	Appleton 2002; Depres et al. 2008; Salzman et al. 2018	Feather et al. 1999; Roberts and Lubowski 2007; Connor et al. 2008		
Voluntary approaches		Khanna and Damon 1999; Khanna 2001; Koehler 2007; Innes and Sam 2008; Borck and Coglianese 2009; Vidovic and Khanna 2012				
Infrastructure investments			Keiser and Shapiro 2018			Cutler and Miller 2005; Watson 2006; Delaney et al. 2011; Alsan and Goldin 2018

Table C.1: Empirical evidence on the impacts of water pollution control policies in developed countries, cont.

		Socioeconomic impacts					Cost-effectiveness	Unintended consequences	
Water pollution control policies		*Water supply behavior*	*Capacity building*	*Capital markets*	*Rural development & income transfers*	*Education and cognitive development*	*Cost-effectiveness*	*Strategic regulatory avoidance*	*Rent-seeking*
Prescriptive policies	*Technology standards*						Carson and Mitchell 1993; Lyon and Farrow 1995; Keiser and Shapiro 2018		
	Performance standards						Carson and Mitchell 1993; Lyon and Farrow 1995; Keiser and Shapiro 2018		
Market-based policies	*Pollution taxes*								
	Tradable permits								
	Information disclosure			Laplante & Lanoie 1994; Hamilton 1995; Dasgupta et al. 2006; Khanna et al. 1998				Greenstone 2003; Gamper-Rabindran and Swoboda 2006; Bennear 2008	
	Payments for ecosystem services				Baylis et al. 2008		Appleton 2002; Connor et al. 2008; Claassen et al. 2008		
Voluntary approaches									
Infrastructure investments							van der Veeren and Tol 2001		

Color code:
Black: causal relations
Blue: descriptive and non-causal empirical studies
Orange: review article
Green: qualitative study

Table C.2: Empirical evidence on the impacts of water pollution control policies in developing countries

		Environmental impacts		Health impacts	
Water pollution control policies		***Reduction in emissions/effluent***	***Ambient water quality improvement***	***Reduced morbidity***	***Reduced premature mortality***
Prescriptive policies	***Technology standards***	Greenstone and Hanna 2014	Greenstone and Hanna 2014		
	Performance standards	Afsah and Makarim 1999; Dasgupta et al. 2000; Duflo et al 2013; Zheng et al. 2018			Do et al. 2018
Market-based policies	***Pollution taxes***	Jiang and McKibbin 2002; Wang and Wheeler 2003; Wang and Wheeler 1999; Wang 2000; He and Zhang 2018; Ebenstein 2012; Blackman 2006 ; Vincent 1993; Vincent et al. 1997; Bluffstone 2003	Ebenstein 2012; Bluffstone 2003		Ebenstein 2012
	Tradable permits				
	Information disclosure	Tietenberg 1998; Blackman et al. 2004; Garcia et al. 2007, 2009; Powers et al. 2011			
	Payments for ecosystem services	Zheng et al. 2013	Moreno-Sanchez et al. 2012, Branca et al. 2011		
Voluntary approaches		Pargal and Wheeler 1996; Blackman et al. 2006; Blackman and Sisto 2006; Blackman et al. 2010; Jiménez 2007; Hu 2007; Blackman et al. 2013	Blackman et al. 2010		
Infrastructure investments				Zhang 2012; Duflo et al. 2015	Galiani et al. 2005; Soares 2007; Geruso and Spears 2018

Table C.2: Empirical evidence on the impacts of water pollution control policies in developing countries, cont.

		Socioeconomic impacts					Cost-effectiveness	Unintended consequences	
Water pollution control policies		***Water supply behavior***	***Capacity building***	***Capital markets***	***Labor market and poverty alleviation***	***Education and cognitive development***	***Cost-effectiveness***	***Strategic regulatory avoidance***	***Rent-seeking***
Prescriptive policies	***Technology standards***								
	Performance standards							Kahn et al. 2015; Chen et al. 2018	Duflo et al. 2013
Market-based policies	***Pollution taxes***		Besley and Persson 2013				Ebenstein 2012	Cai et al. 2016; Lin 2013	
	Tradable permits								
	Information disclosure	Madajewicz et al. 2007; Barnwal et al. 2017; Bennear et al. 2013		Dasgupta et al. 2001a					
	Payments for ecosystem services				Pagiola et al. 2005; Bulte et al. 2008; Zheng et al. 2013; Alix-Garcia et al. 2015, 2018; Branca et al. 2011; Pattanayak et al. 2010, Mussa & Mwakaje 2013; Kosoy et al. 2007		Zheng et al. 2013	Pattanayak et al. 2010	
Voluntary approaches									
Infrastructure investments					Zhang 2012; Spears 2012	Sekhri 2013; Devoto et al. 2012; Zhang and Xu 2016			

Color code:

Black: causal relations

Blue: descriptive and non-causal empirical studies

Orange: review article

Green: qualitative study

Bibliography

AAA (2019), 'Aaa's your driving costs'.
URL: *https://exchange.aaa.com/automotive/driving-costs/#.XLonV2hKhaR*

Abidoye, B., Herriges, J. A. & Tobias, J. L. (2012), 'Controlling for Observed and Unobserved Site Characteristics in RUM Models of Recreation Demand', *American Journal of Agricultural Economics* **94**(5), 1070–1093.

Abidoye, B. O. & Herriges, J. A. (2012), 'Model Uncertainty in Characterizing Recreation Demand', *Environmental and Resource Economics* **53**, 251–277.

Afsah, S. & Makarim, N. (1999), *Program-based pollution control management: the Indonesian PROKASIH program*, The World Bank.

Aizer, A. & Currie, J. (2019), 'Lead and juvenile delinquency: new evidence from linked birth, school, and juvenile detention records', *Review of Economics and Statistics* **101**(4), 575–587.

Aizer, A., Currie, J., Simon, P. & Vivier, P. (2018), 'Do low levels of blood lead reduce children's future test scores?', *American Economic Journal: Applied Economics* **10**(1), 307–41.

Alberini, A. & Segerson, K. (2002), 'Assessing voluntary programs to improve environmental quality', *Environmental and Resource Economics* **22**(1-2), 157–184.

Alix-Garcia, J. M., Sims, K. R., Orozco-Olvera, V. H., Costica, L. E., Fernández Medina, J. D. & Monroy, S. R. (2018), 'Payments for environmental services supported social capital while increasing land management', *Proceedings of the National Academy of Sciences of the United States of America* **115**(27), 7016–7021.

Allaire, M., Wu, H. & Lall, U. (2018), 'National trends in drinking water quality violations', *Proceedings of the National Academy of Sciences* **115**(9), 2078–2083.

Alsan, M. & Goldin, C. (2019), 'Watersheds in child mortality: The role of effective water and sewerage infrastructure, 1880–1920', *Journal of Political Economy* **127**(2), 586–638.

American, S., Journal, E., Policy, E., November, N., Alix-garcia, B. J. M., Sims, K. R. E. & Yanez-pagans, P. (2015), 'American Economic Association Only One Tree from Each Seed ? Environmental Effectiveness and Poverty Alleviation in Mexico ' s Payments for Ecosystem Services Program Author (s): Jennifer M . Alix-Garcia , Katharine R . E . Sims and Patricia Yañez-Pagan', *American Economic Journal: Economic Policy* **7**(4).

American Water Works Association (1996), *Internal corrosion of water distribution systems*, American Water Works Association.

Anderson, T. L. & Libecap, G. D. (2014), *Environmental markets: a property rights approach*, Cambridge University Press.

Andres, L., Briceño, B., Chase, C. & Echenique, J. A. (2017), 'Sanitation and externalities: evidence from early childhood health in rural india', *Journal of Water, Sanitation and Hygiene for Development* **7**(2), 272–289.

Bajari, P., Fruehwirth, J. C., Kim, K. I. & Timmins, C. (2012), 'A rational expectations approach to hedonic price regressions with time-varying unobserved product attributes: The price of pollution', *American Economic Review* **102**(5), 1898–1926.

Baker, J., Bennear, L. & Olmstead, S. (2019), 'Does information disclosure reduce drinking water violations in the united states?'.

Baker, M. (1897), *The Manual of American Water-Works*, The Engineering News Publishing Co.

Banzhaf, H. S. & Chupp, B. A. (2012), 'Fiscal federalism and interjurisdictional externalities: New results and an application to us air pollution', *Journal of Public Economics* **96**(5-6), 449–464.

Banzhaf, H. S., Ma, L. & Timmins, C. (2019), 'Environmental justice: Establishing causal relationships', *Annual Review of Resource Economics* **11**, 377–398.

Barbier, E. B. (2012), 'Progress and challenges in valuing coastal and marine ecosystem services', *Review of Environmental Economics and Policy* **6**(1), 1–19.

Baum, C. F., Schaffer, M. E. & Stillman, S. (2007), 'Enhanced routines for instrumental variables/generalized method of moments estimation and testing', *The Stata Journal* **7**(4), 465–506.

Beaulieu, J. J., Delsontro, T. & Downing, J. (2019), 'Eutrophication will increase methane emissions from lakes and impoundments during the 21st century', *Nature Communications* **10**, 1375.

Beck, M. W., Sherwood, E. T., Henkel, J. R., Dorans, K., Ireland, K. & Varela, P. (2019), 'Assessment of the cumulative effects of restoration activities on water quality in Tampa Bay, Florida', *Working Paper* .

Beharry-Borg, N., Scarpa, R., Beharry-Borg, N. & Scarpa, R. (2010), 'Valuing quality changes in Caribbean coastal waters for heterogeneous beach visitors', *Ecological Economics* **69**(5), 1124–1139.
URL: *https://econpapers.repec.org/RePEc:eee:ecolec:v:69:y:2010:i:5:p:1124-1139*

Bellinger, D. C. (2008), 'Very low lead exposures and children's neurodevelopment', *Current Opinion in Pediatrics* **20**(2), 172–177.

Bellinger, D. C., Stiles, K. M. & Needleman, H. L. (1992), 'Low-level lead exposure, intelligence and academic achievement: A long-term', *Pediatrics* **90**(6), 855.

Benfer, E. A. (2017), 'Contaminated childhood: The chronic lead poisoning of low-income children and communities of color in the united states'.
URL: *https://www.healthaffairs.org/do/10.1377/hblog20170808.061398/full/*

Bennear, L. S. & Coglianese, C. (2005), 'Measuring progress: program evaluation of environmental policies', *Environment: Science and Policy for Sustainable Development* **47**(2), 22–39.

Bento, A., Freedman, M. & Lang, C. (2015), 'Who benefits from environmental regulation? Evidence from the Clean Air Act Amendments', *Review of Economics and Statistics* **97**(3), 610–622.

Blackman, a., Afsah, S. & Ratunanda, D. (2004), 'How do public disclosure pollution control programs work? Evidence from Indonesia', *Human Ecology Review* **11**(3), 235–246.

Blackman, A., Lahiri, B., Pizer, W., Rivera Planter, M. & Muñoz Piña, C. (2010), 'Voluntary environmental regulation in developing countries: Mexico's Clean Industry Program', *Journal of Environmental Economics and Management* **60**(3), 182–192.

Blackman, A., Lyon, T. P. & Sisto, N. (2006), 'Voluntary Environmental Agreements when Regulatory Capacity is Weak', *Comparative Economic Studies* **48**(4), 682–702.

Blackman, a., Morgenstern, R. & Topping, E. (2006), 'Institutional analysis of Colombia's autonomous regional corporations (CARs)', (June).

Blackman, A. & Sisto, N. (2006), 'Voluntary environmental regulation in developing countries: A Mexican case study', *Natural Resources Journal* **46**(4), 1005–1042.

Blackman, A., Uribe, E., van Hoof, B. & Lyon, T. P. (2013), 'Voluntary environmental agreements in developing countries: The Colombian experience', *Policy Sciences* **46**(4), 335–385.

Blackwell, M., Honaker, J. & King, G. (2017), 'A unified approach to measurement error and missing data: overview and applications', *Sociological Methods & Research* **46**(3), 303–341.

Bockstael, N. E., Hanemann, M. W. & Kling, C. L. (1987), 'Estimating the Value of Water Quality Improvements in a Recreational Demand Framework', *Water Resources Research* **23**(5), 951–960.

Bockstael, N. E., McConnell, K. E. & Strand, I. E. (1989), 'Measuring the Benefits of Improvements in Water Quality: The Chesapeake Bay', *Marine Resource Economics* **6**(1), 280–290.

Borck, J. C. & Coglianese, C. (2009), 'Voluntary environmental programs: Assessing their effectiveness', *Annual Review of Environment and Resources* **34**, 305–324.

Borden, J. C., Borden, S. & Mistry, P. (2016), Crowdsourcing water quality data: a conceptual framework, Technical report, The World Bank.

Boyd, J. (2003), 'Water Pollution Taxes: A Good Idea Doomed to Failure?', *Public Finance and Management* **3**(1), 34–66.

Boyle, K. J. & Bouchard, R. (2003), 'Water Quality Effects On Property Prices In Northern New England', *Lake Line* **23**(3), 24–27.

Boyle, K. J., Poor, J. P. & Taylor, L. O. (1999), 'Estimating the Demand for Protecting Freshwater Lakes from Eutrophication', *American Journal of Agricultural Economics* **81**(5), 1118–1122.

Branca, G., Lipper, L., Neves, B., Lopa, D. & Mwanyoka, I. (2011), 'Payments for watershed services supporting sustainable agricultural development in Tanzania', *Journal of Environment and Development* **20**(3), 278–302.

Bressers, H. T. A. (1988), 'A comparison of the effectiveness of incentives and directives: the case of dutch water quality policy', *Review of Policy Research* **7**(3), 500–518.

Bulte, E. H., Lipper, L., Stringer, R. & Zilberman, D. (2008), 'Payments for ecosystem services and poverty reduction: Concepts, issues, and empirical perspectives', *Environment and Development Economics* **13**(3), 245–254.

Cai, H., Chen, Y. & Gong, Q. (2016), 'Polluting thy neighbor: Unintended consequences of China's pollution reduction mandates', *Journal of Environmental Economics and Management* **76**.

Cameron, A. C. & Miller, D. L. (2015), 'A practitioners guide to cluster-robust inference', *Journal of Human Resources* **50**(2), 317–372.

Campbell, S. & Wessel, D. (2021), 'What would it cost to replace all the nation's lead water pipes?'.
URL: *https://www.brookings.edu/blog/up-front/2021/05/13/what-would-it-cost-to-replace-all-the-nations-lead-water-pipes/*

Carnevale, A. P., Cheah, B. & Rose, S. J. (2011), 'The college pay off'.

CDC (2004).
URL: *https://www.cdc.gov/nceh/lead/acclpp/meetingminutes/lessthan10mtgmar04.pdf*

Cesario, F. J. (1976), 'Value of time in recreation benefit studies', *Land Economics* **52**(1), 32–41.
URL: *http://www.jstor.org/stable/3144984*

Chen, A., Cai, B., Dietrich, K. N., Radcliffe, J. & Rogan, W. J. (2007), 'Lead Exposure, IQ, and Behavior in Urban 5- to 7-Year-Olds: Does Lead Affect Behavior Only by Lowering IQ?', *Pediatrics* **119**(3), e650–e658.

Chen, Z., Kahn, M. E., Liu, Y. & Wang, Z. (2018), 'The consequences of spatially differentiated water pollution regulation in China', *Journal of Environmental Economics and Management* **88**, 468–485.
URL: *http://dx.doi.org/10.1016/j.jeem.2018.01.010*

Chetty, R., Hendren, N., Lin, F., Majerovitz, J. & Scuderi, B. (2016), 'Childhood environment and gender gaps in adulthood', *American Economic Review* **106**(5), 282–88.

Chiang, J. Q. W. (2016), Does decentralization work differently when externality differs ? Evidence from the Chinese water sector.

Choe, K., Whittington, D. & Lauria, D. T. (1996), 'The economic benefits of surface water quality improvements in developing countries: a case study of davao, philippines', *Land Economics* pp. 519–537.

Clay, K., Troesken, W. & Haines, M. (2014*a*), 'Lead and mortality', *Review of Economics and Statistics* **96**(3), 458–470.

Clay, K., Troesken, W. & Haines, M. (2014*b*), 'Lead and Mortality', *Review of Economics and Statistics* **96**(3), 458–470.

Clean Water Network of Florida (2008), The Gulf of Mexico – Florida's Toilet, Technical Report June.

Coase, R. H. (1960), The problem of social cost, *in* 'Classic papers in natural resource economics', Springer, pp. 87–137.

Columbia Water Center (2016).
URL: *http://water.columbia.edu/files/2016/03/Developing-a-Water-Road-Map-Whitepaper.pdf*

Committee on Environmental Health and others (2005), 'Lead exposure in children: prevention, detection, and management', *Pediatrics* **116**(4), 1036–1046.

Cunningham, S. (2021), *Causal Inference*, Yale University Press.

Currie, J., Graff Zivin, J., Meckel, K., Neidell, M. & Schlenker, W. (2013), 'Something in the water: Contaminated drinking water and infant health', *Canadian Journal of Economics/Revue canadienne d'économique* **46**(3), 791–810.

Currie, J., Zivin, J. G., Mullins, J. & Neidell, M. (2014), 'What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?', *Annual Review of Resource Economics* **6**(1), 217–247.

Cutler, D. & Miller, G. (2005), 'The role of public health improvements in health advances: The twentieth-century United States', *Demography* **42**(1), 1–22.

Dakkak, A. (2018), 'Water pollution worries in developing world'.

d'Arge, R. C. & Shogren, J. F. (1989), 'Okoboji Experiment: Comparing Non-Market Valuation Techniques In An Unusually Well-Defined Market For Water Quality', *Ecological Economics* **1**(3), 251–259.

Dasgupta, S., Hettige, H. & Wheeler, D. (2000), 'What improves environmental compliance? Evidence from Mexican industry', *Journal of Environmental Economics and Management* **39**(1), 39–66.

Dasgupta, S., Laplante, B., Mamingi, N. & Wang, H. (2001), 'Inspections, pollution prices, and environmental performance: Evidence from China', *Ecological Economics* **36**(3), 487–498.

Dasgupta, S., Wheeler, D. & Wang, H. (2006), 'Disclosure strategies for pollution control', *International yearbook of environmental and resource economics* **2007**, 93–119.

Day, B. & Mourato, S. (2002), *Valuing river water quality in China*, Edward Elgar.

Delaney, L., McGovern, M. & Smith, J. P. (2011), 'From angela's ashes to the celtic tiger: Early life conditions and adult health in ireland', *Journal of Health Economics* **30**(1), 1–10.

Dell, M., Jones, B. F. & Olken, B. A. (2014), 'What do we learn from the weather? the new climate-economy literature', *Journal of Economic Literature* **52**(3), 740–98.

DeSantis, M. K., Triantafyllidou, S., Schock, M. R. & Lytle, D. A. (2018), 'Mineralogical evidence of galvanic corrosion in drinking water lead pipe joints', *Environmental science & technology* **52**(6), 3365–3374.

Deschenes, O. & Meng, K. C. (2018), Quasi-experimental methods in environmental economics: Opportunities and challenges, *in* 'Handbook of Environmental Economics', Vol. 4, Elsevier, pp. 285–332.

Do, Q.-T., Joshi, S. & Stolper, S. (2018), 'Can environmental policy reduce infant mortality? evidence from the ganga pollution cases', *Journal of Development Economics* **133**, 306–325.

Duflo, E., Greenstone, M., Guiteras, R. & Clasen, T. (2015), Toilets Can Work: Short and Medium Run Health Impacts of Addressing Complementarities and Externalities in Water and Sanitation.

Duflo, E., Greenstone, M., Pande, R. & Ryan, N. (2013), 'Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India', *The Quarterly Journal of Economics* (April), 1499–1545.

Ebenstein, A. (2012), 'The Consequences of Industrialization : Evidence From Water Pollution and Digestive Cancers in China', *The Review of Economics and Statistics* **94**(February), 186–201.

Edwards, M. & Triantafyllidou, S. (2007), 'Chloride-to-sulfate mass ratio and lead leaching to water', *Journal-American Water Works Association* **99**(7), 96–109.

Edwards, M., Triantafyllidou, S. & Best, D. (2009), 'Elevated blood lead in young children due to lead-contaminated drinking water: Washington, dc, 2001- 2004', *Environmental science & technology* **43**(5), 1618–1623.

Egan, K. J., Herriges, J. A., Kling, C. L. & Downing, J. A. (2009), 'Valuing

Water Quality as a Function of Water Quality Measures', *American Journal of Agricultural Economics* **91**(1), 106–123.

Environmental Defense Fund (2020).
URL: *https://www.edf.org/sites/default/files/u4296/LeadPipe_EnvironJustice_AU%20and%20EDF%20Report.pdf*

Epp, J. D. & Al-Ani, K. S. (1979), 'The Effect of Water Quality on Rural Nonfarm Residential Property Values', *American Journal of Agricultural Economics* **61**(3), A1–A16.

Feather, P., Hellerstein, D. & Hansen, L. (1999), 'Economic valuation of environmental benefits and the targeting of conservation programs: The case of the crp', *Economic Research Service AER* **778**.

Fedinick, K. P., Wu, M., Panditharatne, M. & Olson, E. D. (2017).
URL: *https://www.nrdc.org/sites/default/files/threats-on-tap-water-infrastructure-protections-report.pdf*

Feigenbaum, J. J. & Muller, C. (2016), 'Lead exposure and violent crime in the early twentieth century', *Explorations in Economic History* **62**, 51–86.

Ferraro, P. J. (2008), 'Asymmetric information and contract design for payments for environmental services', *Ecological economics* **65**(4), 810–821.

Ferrie, J. P., Rolf, K. & Troesken, W. (2012), 'Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among World War Two U.S. Army enlistees', *Economics and Human Biology* **10**(1), 98–111.

Fisher-vanden, K. & Olmstead, S. (2013), 'Moving Pollution Trading from Air to Water: Potential, Problems, and Prognosis', *Journal of Economic Perspectives* **27**(1), 147–171.

Florida Fish and Wildlife Conservation Commission (2003), Florida Seagrass Manager's Toolkit, Technical Report September.

Freeman, A. M., Herriges, J. A. & Kling, C. L. (2014), *The Measurement of Environmental and Resource Values: Theory and Methods*, Routledge.

Gallicchio, L., Scherer, R. W. & Sexton, M. (2002), 'Influence of nutrient intake on blood lead levels of young children at risk for lead poisoning.', *Environmental Health Perspectives* **110**(12), A767–A772.

García, J. H., Afsah, S. & Sterner, T. (2009), 'Which firms are more sensitive to public disclosure schemes for pollution control? Evidence from Indonesia's PROPER program', *Environmental and Resource Economics* **42**(2), 151–168.

García, J. H., Sterner, T. & Afsah, S. (2007), 'Public disclosure of industrial pollution: The PROPER approach for Indonesia?', *Environment and Development Economics* **12**(6), 739–756.

Garg, T., Hamilton, S. E., Hochard, J. P., Kresch, E. P. & Talbot, J. (2018), '(not so) gently down the stream: River pollution and health in indonesia', *Journal of Environmental Economics and Management* **92**, 35–53.

Ghanem, D. & Zhang, J. (2014), "effortless perfection:'do chinese cities manipulate air pollution data?', *Journal of Environmental Economics and Management* **68**(2), 203–225.

Gibbs, J. P., Halstead, J. M., Boyle, K. J. & Huang, J. C. (2002), 'An hedonic analysis of the effects of lake water clarity on New Hampshire lakefront properties', *Agricultural and Resource Economics Review* **31**(1), 39–46.

Greening, H., Janicki, A., Sherwood, E. T., Pribble, R. & Johansson, J. O. R. (2014), 'Ecosystem responses to long-term nutrient management in an urban estuary: Tampa bay, florida, usa', *Estuarine, Coastal and Shelf Science* **151**, A1–A16.

Greenstone, M. & Gallagher, J. (2008), 'Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program', *Quarterly Journal of Economics* (August), 951–1003.

Greenstone, M. & Hanna, R. (2014), 'Environmental Regulations, Air and Water Pollution, and Infant Mortality in India', *American Economic Review* **104**(10), 3038–3072.

Greenstone, M. & Jack, B. K. (2015), 'Envirodevonomics: A research agenda for an emerging field', *Journal of Economic Literature* **53**(1), 5–42.

Grooms, K. K. (2016), 'Does water quality improve when a safe drinking water act violation is issued? a study of the effectiveness of the sdwa in california', *The BE Journal of Economic Analysis & Policy* **16**(1), 1–23.

Grönqvist, H., Nilsson, J. P. & Robling, P.-O. (2020), 'Understanding how low levels of early lead exposure affect children's life trajectories', *Journal of Political Economy* **128**(9), 3376–3433.

Guignet, D., Griffiths, C., Klemick, H. & Walsh, P. J. (2017), 'The Implicit Price of Aquatic Grasses', *Marine Resource Economics* **32**(1), 21–41.

Haab, T. C. & McConnell, K. E. (2002), *Valuing Environmental and Natural Resources*, Vol. 8.

Hagerty, N. (2018), 'The costs of industrial water pollution to agriculture in india'.

Hamilton, J. T. (1995), 'Pollution as news: Media and stock market reactions to the toxics release inventory data', *Journal of environmental economics and management* **28**(1), 98–113.

Hanna-Attisha, M., LaChance, J., Sadler, R. C. & Schnepp, A. C. (2016), 'Elevated blood lead levels in children associated with the flint drinking water crisis: A spatial analysis of risk and public health response', *American Journal of Public Health* **106**(2), 283–290.

He, G., Wang, S. & Zhang, B. (2020), 'Watering down environmental regulation in china', *The Quarterly Journal of Economics* **135**(4), 2135–2185.

He, P. & Zhang, B. (2018), ‘Environmental Tax, Polluting Plants’ Strategies and Effectiveness: Evidence from China’, *Journal of Policy Analysis and Management* **37**(3), 493–520.

Heckman, J. J. & Karapakula, G. (2019), The perry preschoolers at late midlife: A study in design-specific inference, Technical report, National Bureau of Economic Research.

Hollingsworth, A. & Rudik, I. (2020), ‘The social cost of leaded gasoline: Evidence from regulatory exemptions’.

Horsch, E. J. & Lewis, D. J. (2009), ‘The effects of aquatic invasive species on property values: evidence from a quasi-experiment.’, *Land Economics* **85**(3), 391–409.

Hu, Y. (2007), ‘Implementation of voluntary agreements for energy efficiency in China’, *Energy Policy* **35**(11), 5541–5548.

Innes, R. & Sam, A. G. (2008), ‘Voluntary pollution reductions and the enforcement of environmental law: An empirical study of the 33/50 program’, *The Journal of Law and Economics* **51**(2), 271–296.

Isen, A., Rossin-Slater, M. & Walker, W. R. (2017), ‘Every Breath You Take—Every Dollar You’ll Make: The Long-Term Consequences of the Clean Air Act of 1970’, *Journal of Political Economy* **125**(3), 848–902.

Jack, B. K., Kousky, C. & Sims, K. R. (2008), ‘Designing payments for ecosystem services: Lessons from previous experience with incentive-based mechanisms', *Proceedings of the National Academy of Sciences* **105**(28), 9465–9470.

Jacobs, D. E., Clickner, R. P., Zhou, J. Y., Viet, S. M., Marker, D. A., Rogers, J. W., Zeldin, D. C., Broene, P. & Friedman, W. (2002), 'The prevalence of lead-based paint hazards in us housing.', *Environmental health perspectives* **110**(10), A599–A606.

Jiang, T. & McKibbin, W. J. (2002), 'Assessment of China ' s pollution levy system : an equilibrium pollution approach', *Environment and Development Economics* **7**, 75–105.

Jiménez, O. (2007), 'Voluntary agreements in environmental policy: an empirical evaluation for the Chilean case', *Journal of Cleaner Production* **15**(7), 620–637.

Johansson, J. O. R. (2016), Seagrass Transect Monitoring in Tampa Bay A Summary of Findings from 1997 through 2015 Prepared for the Tampa Bay Estuary Program, Technical report.

Kahn, M. E., Li, P. & Zhao, D. (2015), 'Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives', *American Economic Journal: Economic Policy* **7**(4), 223–242.

Keiser, D. A., Kling, C. L. & Shapiro, J. S. (2019), 'The low but uncertain measured benefits of US water quality policy', *Proceedings of the National Academy of Sciences* **116**(12), 5262–5269.

Keiser, D. A. & Shapiro, J. S. (2019*a*), 'Burning Waters to Crystal Springs? U.S. Water Pollution Regulation over the Last Half Century', *NBER Working Paper* .

Keiser, D. A. & Shapiro, J. S. (2019*b*), 'Consequences of the Clean Water Act and the demand for water quality', *Quarterly Journal of Economics* **134**(1), 349–396.

Keiser, D. A. & Shapiro, J. S. (2019*c*), 'Consequences of the clean water act and the demand for water quality', *The Quarterly Journal of Economics* **134**(1), 349–396.

Khanna, M. & Damon, L. A. (1999), 'Epa's voluntary 33/50 program: Impact on toxic releases and economic performance of firms', *Journal of environmental economics and management* **37**(1), 1–25.

Khanna, M., Quimio, W. R. H. & Bojilova, D. (1998), 'Toxics release information: A policy tool for environmental protection', *Journal of environmental economics and management* **36**(3), 243–266.

Kim, Y., Kling, C. L. & Zhao, J. (2015), 'Understanding behavioral explanations of the wtp-wta divergence through a neoclassical lens: implications for environmental policy', *Annu. Rev. Resour. Econ.* **7**(1), 169–187.

Konishi, Y., Coggins, J. S. & Wang, B. (2015), 'Water-quality trading: Can we get the prices of pollution right?', *Water Resources Research* **51**(5), 3126–3144.

Kosoy, N., Martinez-Tuna, M., Muradian, R. & Martinez-Alier, J. (2007), 'Payments for environmental services in watersheds: Insights from a comparative study of three cases in Central America', *Ecological Economics* **61**(2-3), 446–455.

Kremer, M., Leino, J., Miguel, E. & Zwane, A. P. (2011), 'Spring cleaning: Rural water impacts, valuation, and property rights institutions', *Quarterly Journal of Economics* **126**(1), 145–205.

Kresch, E. P., Lipscomb, M. & Schechter, L. (2020), 'Externalities and spillovers from sanitation and waste management in urban and rural neighborhoods', *Applied Economic Perspectives and Policy* **42**(3), 395–420.

Kuwayama, Y., Olmstead, S. M., Wietelman, D. C. & Zheng, J. (2020), 'Trends in nutrient-related pollution as a source of potential water quality damages: A case study of texas, usa.', *The Science of the total environment* **724**, 137962.

Lanphear, B. P., Dietrich, K., Auinger, P. & Cox, C. (2000), 'Cognitive deficits associated with blood lead concentrations $<$10 microg/dL in US children and Adolescents', *Public health reports* **115**(6), 521–9.
URL: *http://www.ncbi.nlm.nih.gov/pubmed/11354334%5Cnhttp://www.pubmedcentral.nih.gov*

Lanphear, B. P. & Roghmann, K. J. (1997), 'Pathways of lead exposure in urban children', *Environmental Research* **74**(1), 67–73.

Laplante, B. & Lanoie, P. (1994), 'The market response to environmental incidents in canada: a theoretical and empirical analysis', *Southern Economic Journal* pp. 657–672.

Leggett, C. G. & Bockstael, N. E. (2000), 'Evidence of the Effects of Water Quality on Residential Land Prices', *Journal of Environmental Economics and Management* **39**, 121–144.

Lin, L. (2013), 'Enforcement of pollution levies in China', *Journal of Public Economics* **98**, 32–43.
URL: *http://dx.doi.org/10.1016/j.jpubeco.2012.11.004*

Lipscomb, M. & Mobarak, A. M. (2017), 'Decentralization and pollution spillovers: Evidence from the re-drawing of county borders in Brazil', *Review of Economic Studies* **84**(1), 464–502.

Luxen, D. & Vetter, C. (2011), Real-time routing with OpenStreetMap data, *in* 'Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems', GIS '11, ACM, New York, NY, USA, pp. 513–516.

Marcus, M. (2020), 'Testing the water: Drinking water quality, public notification, and child outcomes'.

Massey, D. M., Newbold, S. C. & Gentner, B. (2006), 'Valuing water quality changes using a bioeconomic model of a coastal recreational fishery', *Journal of Environmental Economics and Management* **52**(1), 482–500.

Masten, S. J., Davies, S. H. & Mcelmurry, S. P. (2016), 'Flint water crisis: what happened and why?', *Journal-American Water Works Association* **108**(12), 22–34.

Meeks, R. C. (2017), 'Water works: The economic impact of water infrastructure', *Journal of Human Resources* **52**(4), 1119–1153.

Mendelsohn, R., Hellerstein, D., Huguenin, M., Unsworth, R. & Brazee, R. (1992), 'Measuring hazardous waste damages with panel models', *Journal of Environmental Economics and Management* **22**(3), 259–271.

Mishra, P. P. (2017), 'The benefits of improving urban lakes in mega cities: a revealed and stated preference approach applied to the hussain sagar in hyderabad, india', *Environment and Development Economics* **22**(4), 447–469.

Moeltner, K. & von Haefen, R. (2011), 'Microeconometric Strategies for Dealing with Unobservables and Endogenous Variables in Recreation Demand Models', *Annual Review of Resource Economics* **3**(1), 375–396.

Montero, J.-P., Sanchez, J. M. & Katz, R. (2002), 'A market-based environmental policy experiment in chile', *The Journal of Law and Economics* **45**(1), 267–287.

Moreno-Sanchez, R., Maldonado, J. H., Wunder, S. & Borda-Almanza, C. (2012), 'Heterogeneous users and willingness to pay in an ongoing payment for watershed protection initiative in the Colombian Andes', *Ecological Economics* **75**, 126–134.

URL: *http://dx.doi.org/10.1016/j.ecolecon.2012.01.009*

Morrison, G. & Greening, H. (2006), Chapter 5 . Water Quality, Technical report, U.S.Department of the Interior, U.S. Geological Survey,Tampa Bay Estuary Program.

Muehlenbachs, L., Spiller, E. & Timmins, C. (2015), 'The Housing Market Impacts of Shale Gas Development', *American Economic Review* **105**(12), 3633–3659.

Mullen, J. K. & Menz, F. C. (1985), 'The Effect of Acidification Damages on the Economic Value of the Adirondack Fishery to New York Anglers', *American Journal of Agricultural Economics* **67**(1), 112–119.

Mussa, K. R. & Mwakaje, A. G. (2013), 'The impact of equitable payment for watershed services scheme on livelihoods in tanzania: The case of uluguru mountains', *International Journal of Development and Sustainability* **2**(2), 1031–1051.

Netusil, N. R., Kincaid, M. & Chang, H. (2014), 'Valuing water quality in urban watersheds: A comparative analysis of Johnson Creek, Oregon, and Burnt Bridge Creek, Washington', *Water Resources Research* **50**(5), 4254–4268.

New Hampshire Department of Environmental Services (2010), 'Sodium and chloride in drinking water'.
URL: *https://www.cambridgema.gov/-/media/Files/waterdepartment/Distribution/ChloridesE*

Newell, R. & Rogers, K. (2003), The US Experience with the Phasedown of Lead in Gasoline.

NOAA Fisheries (2008), Marine Recreational Information Program: Implementation Plan, Technical Report October.

NOAA Fisheries (2013), Marine Recreational Information Program Data User Handbook, Technical report.

Oates, W. E. & Schwab, R. M. (1988), 'Economic competition among jurisdictions: efficiency enhancing or distortion inducing?', *Journal of public economics* **35**(3), 333–354.

Olson, E. & Fedinick, K. P. (2016).
URL: *https://www.nrdc.org/sites/default/files/whats-in-your-water-flint-beyond-report.pdf*

OpenStreetMap contributors (2018), 'Florida parks map retrieved from `https://planet.osm.org`', `https://www.openstreetmap.org`.

Ouellet, V., St-Hilaire, A., Dugdale, S. J., Hannah, D. M., Krause, S. & Proulx-Ouellet, S. (2020), 'River temperature research and practice: Recent

challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems', *Science of The Total Environment* **736**, 139679.
URL: *https://www.sciencedirect.com/science/article/pii/S0048969720331995*

Pagiola, S., Arcenas, A. & Platais, G. (2005), 'Can Payments for Environmental Services help reduce poverty? An exploration of the issues and the evidence to date from Latin America', *World Development* **33**(2 SPEC. ISS.), 237–253.

Palmquist, R. B. (1982), 'Measuring environmental effects on property values without hedonic regressions', *Journal of Urban Economics* **11**(3), 333–347.

Panno, S. V., Hackley, K. C., Hwang, H., Greenberg, S. E., Krapac, I., Landsberger, S. & O'kelly, D. (2006), 'Characterization and identification of na-cl sources in ground water', *Groundwater* **44**(2), 176–187.

Pattanayak, S. K., Wunder, S. & Ferraro, P. J. (2010), 'Show me the money: Do payments supply environmental services in developing countries?', *Review of Environmental Economics and Policy* **4**(2), 254–274.

Pelley, J. (2018), 'Treatment for lead in drinking water is evolving. will the u.s. epa catch up?'.
URL: *https://cen.acs.org/environment/water/Treatment-lead-drinking-water-evolving/96/i47*

Phaneuf, D. (2013), 'Heterogeneity in environmental demand', *Annual Reviews of Resource Economics* **5**, 227–244.

Phaneuf, D. J. (2002), 'A random utility model for total maximum daily loads: Estimating the benefits of watershed-based ambient water quality improvements', *Water Resources Research* **38**(11), 1254.

Phaneuf, D. J., Kling, C. L. & Herriges, J. A. (2000), 'Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand', *Review of Economics and Statistics* **82**(1), 83–92.

Phaneuf, D. J., Smith, V., Palmquist, R. & Pope, J. (2008), 'Integrating property value and local recreation models to value ecosystem services in urban watersheds', *Land Economics* **84**(3), 361–381.

Pieper, K. J., Martin, R., Tang, M., Walters, L., Parks, J., Roy, S., Devine, C. & Edwards, M. A. (2018), 'Evaluating water lead levels during the flint water crisis', *Environmental science & technology* **52**(15), 8124–8132.

Pieper, K. J., Tang, M. & Edwards, M. A. (2017), 'Flint water crisis caused by interrupted corrosion control: investigating "ground zero" home', *Environmental Science & Technology* **51**(4), 2007–2014.

Poor, P. J., Boyle, K. J., Taylor, L. O. & Bouchard, R. (2001), 'Objective Versus Subjective Measures of Water Clarity in Hedonic Property Value Models', *Land Economics* **77**(4), 482–493.

Poor, P. J., Pessagno, K. L. & Paul, R. W. (2007), 'Exploring the hedonic value of ambient water quality: A local watershed-based study', *Ecological Economics* **60**(4), 797–806.

Powers, N., Blackman, A., Lyon, T. P. & Narain, U. (2011), 'Does Disclosure Reduce Pollution? Evidence from India's Green Rating Project', *Environmental and Resource Economics* **50**(1), 131–155.

Rabotyagov, S. S., Kling, C. L., Gassman, P. W., Rabalais, N. N. & Turner, R. E. (2014), 'The economics of dead zones: Causes, impacts, policy challenges, and a model of the Gulf of Mexico Hypoxic Zone', *Review of Environmental Economics and Policy* **8**(1), 58–79.

Raymond, J. & Brown, M. (2017), Childhood blood lead level in children age, Technical Report 3.

Reyes, J. W. (2015*a*), 'Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents', *Economic Inquiry* **53**(3), 1580–1605.

Reyes, J. W. (2015*b*), 'Lead Policy and Academic Performance: Insights from Massachusetts', *Harvard Educational Review* **85**(1), 75–107.

Roberts, M. J. & Lubowski, R. N. (2007), 'Enduring impacts of land retirement policies: Evidence from the conservation reserve program', *Land Economics* **83**(4), 516–538.

Rosen, S. (1974), 'Hedonic Prices and Implicit Markets: Production Differentiation in Pure Competition', *Journal of Political Economy* **82**(1), 34–55.

Roy, S. & Edwards, M. A. (2019), 'Preventing another lead (pb) in drinking water crisis: Lessons from the washington dc and flint mi contamination events', *Current Opinion in Environmental Science & Health* **7**, 34–44.

Rubin, D. B. (1987), *Multiple imputation for nonresponse in surveys*, John Wiley & Sons.

Russsell, M. & Greening, H. (2019), 'Estimating benefits in a recovering estuary: Tampa Bay, Florida', *Estuaries and Coasts* **38**(Suppl 1), S9–S18.

Salzman, J., Bennett, G., Carroll, N., Goldstein, A. & Jenkins, M. (2018), 'The global status and trends of payments for ecosystem services', *Nature Sustainability* **1**(3), 136–144.

Sampson, R. J. & Winter, A. S. (2016), 'The racial ecology of lead poisoning: Toxic inequality in chicago neighborhoods, 1995-2013', *Du Bois Review: Social Science Research on Race* **13**(2), 261–283.

Sampson, R. & Winter, A. (2018), 'Poisoned development: assessing childhood lead exposure as a cause of crime in a birth cohort followed through adolescence', *Criminology* **56**(2), 269–301.

Santucci, R. J. & Scully, J. R. (2020), 'The pervasive threat of lead (pb) in drinking water: Unmasking and pursuing scientific factors that govern lead release', *Proceedings of the National Academy of Sciences* **117**(38), 23211–23218.

Schaeffer, B. A., Hagy, J. D., Conmy, R. N., Lehrter, J. C. & Stumpf, R. P. (2012), 'An approach to developing numeric water quality criteria for coastal waters using the seawifs satellite data record', *Environmental science & technology* **46**(2), 916–922.

Schlenker, W. & Roberts, M. J. (2009), 'Nonlinear temperature effects indicate severe damages to us crop yields under climate change', *Proceedings of the National Academy of sciences* **106**(37), 15594–15598.

Schwartz, J. (1994), 'Low-Level Lead Exposure and Children's IQ: A Meta analysis and Search for a Threshold'.

Sherwood, E. T., Greening, H. S., Janicki, A. J. & Karlen, D. J. (2016), 'Tampa Bay estuary: Monitoring long-term recovery through regional partnerships', *Regional Studies in Marine Science* **4**, 1–11.

Shimshack, J. P. (2014), 'The Economics of Environmental Monitoring and Enforcement', *The Annual Review of Resource Economics* **6**, 339–360.

Shortle, J. S. & Horan, R. D. (2001), 'The economics of nonpoint pollution control', *Journal of economic surveys* **15**(3), 255–289.

Sigman, H. (2005), 'Transboundary spillovers and decentralization of environmental policies', *Journal of Environmental Economics and Management* **50**(1), 82–101.

Smith, M. D., Oglend, A., Kirkpatrick, A. J., Asche, F., Bennear, L. S., Craig, J. K. & Nance, J. M. (2017), 'Seafood prices reveal impacts of a major ecological disturbance', *Proceedings of the National Academy of Sciences* **114**(7), 1512–1517.

Smith, V. K., Desvousges, W. H. & Fisher, A. (1986), 'A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits', *American Journal of Agricultural Economics* **68**(2), 280–290.

Soares, R. R. (2007), 'Health and the evolution of welfare across brazilian municipalities', *Journal of Development Economics* **84**(2), 590–608.

Sorensen, L. C., Fox, A. M., Jung, H. & Martin, E. G. (2019), 'Lead exposure and academic achievement: evidence from childhood lead poisoning prevention efforts', *Journal of Population Economics* **32**(1), 179–218.

State Council (2015), 'Water pollution prevention and control action plan'.

Steele, M. & Aitkenhead-Peterson, J. (2011), 'Long-term sodium and chloride surface water exports from the dallas/fort worth region', *Science of the total environment* **409**(16), 3021–3032.

Steinnes, N. D. (1992), 'Measuring the economic value of water quality: The case of lakeshore land', *Annals of Regional Science* **26**(2), 171–176.

Stets, E. G., Lee, C. J., Lytle, D. A. & Schock, M. R. (2018), 'Increasing chloride in rivers of the conterminous us and linkages to potential corrosivity and lead action level exceedances in drinking water', *Science of the Total Environment* **613**, 1498–1509.

Stets, E. G., Sprague, L. A., Oelsner, G. P., Johnson, H. M., Murphy, J. C., Ryberg, K., Vecchia, A. V., Zuellig, R. E., Falcone, J. A. & Riskin, M. L. (2020), 'Landscape drivers of dynamic change in water quality of us rivers', *Environmental Science & Technology* **54**(7), 4336–4343.

Stock, J. H., Wright, J. H. & Yogo, M. (2002), 'A survey of weak instruments and weak identification in generalized method of moments', *Journal of Business & Economic Statistics* **20**(4), 518–529.

Stock, J. & Yogo, M. (2005), *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press, New York, pp. 80–108.

Tampa Bay Estuary Program & Tampa Bay Regional Planning Council (2014), Economic Valuation of Tampa Bay, Technical report, Tampa Bay Estuary Program, Tampa Bay Regional Planning Council.

Texas Education Agency (2015).

Texas Natural Resources Information System (2014), 'Watershed boundaries 2014'.
URL: *https://data.tnris.org/*

Tiebout, C. M. (1956), 'A pure theory of local expenditures', *Journal of political economy* **64**(5), 416–424.

Tietenberg, T. (1998), 'Disclosure Strategies for Pollution Control', *Environmental and Resource Economics* **11**(3-4), 587–602.

Torrice, M. (2020), 'How lead ended up in flint's tap water'.
URL: *https://cen.acs.org/articles/94/i7/Lead-Ended-Flints-Tap-Water.html*

University of South Florida Water Institute (2017), 'Water Altas'.
URL: *http://www.wateratlas.usf.edu/atlases.aspx*

U.S. Bureau of Labor Statistics (2018), 'Occupational Employment Statistics'.
URL: *https://www.bls.gov/oes/home.htm*

U.S. Bureau of Labor Statistics (2019), 'Median weekly earnings $606 for high school dropouts, $1,559 for advanced degree holders'.
URL: *https://www.bls.gov/opub/ted/2019/median-weekly-earnings-606-for-high-school-dropouts-1559-for-advanced-degree-holders.htm*

US Census Bureau (2020), 'Tiger/line shapefiles'.
URL: *https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html*

U.S. Department of Housing and Urban Development (2015), 'Comprehensive housing market analysis: Tampa-St. Petersburg-Clearwater, Florida'.
URL: *https://www.huduser.gov/portal/publications/pdf/TampaFL_comp_15.pdf*

U.S. Environmental Protection Agency (1994), 'Water Quality Criteria', *Water Quality Standards Hand Book* (August), 1–58.

U.S. Environmental Protection Agency (2001), Parameters of water quality, Technical report.

U.S. Environmental Protection Agency (2003), *Drinking water advisory: consumer acceptability advice and health effects analysis on sodium*, U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division.

US Environmental Protection Agency (2016), 'Lead and copper rule revisions white paper'.

U.S. Environmental Protection Agency (2017), Proposed Modeling Approaches for a Health- Based Benchmark for Lead in Drinking Water, Technical report.
URL: *https://www.epa.gov/sites/production/files/2017-01/documents/report_proposed_modeling_approaches_for_a_health_based_benchmark_for_*

U.S. EPA (2000), Lead and Copper Rule : Summary of Revisions, Technical Report April.

US EPA (2019), 'Drinking water dashboard help | echo | us epa'.
URL: *https://echo.epa.gov/help/drinking-water-dashboard-help*

US EPA (2021*a*), 'How lead gets into drinking water'.
URL: *https://www.epa.gov/ground-water-and-drinking-water/basic-information-about-lead-drinking-water#getinto*

US EPA (2021*b*), *national primary drinking water regulations: lead and copper rule revisions*, Vol. 86 FR 4198.

USDA (2018), 'Conservation reserve program monthly summary. october.'. **URL:** *https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/Conservation/PDF/OCT2018%20Summary.pdf*

USEPA (2012), 'Small drinking water systems variances'.

Vidovic, M. & Khanna, N. (2012), 'Is voluntary pollution abatement in the absence of a carrot or stick effective? evidence from facility participation in the epa's 33/50 program', *Environmental and Resource Economics* **52**(3), 369–393.

Vincent, J. R., Ali, R. M., Tan, C., Yahaya, J., Rahim, K., Ghee, L., Meyer, A., Othman, M., Sivalingam, G. et al. (1997), *Environment and development in a resource-rich economy: Malaysia under the New Economic Policy.*, Harvard Institute for International Development, Harvard University.

von Haefen, R. H. (2003), 'Incorporating observed choice into the construction of welfare measures from random utility models', *Journal of Environmental Economics and Management* **45**, 145–165.

Walls, M., Kousky, C. & Chu, Z. (2015), 'Is What You See What You Get? The Value of Natural Landscape Views', *Land Economics* **91**(1), 1–19.

Walsh, P., Griffiths, C., Guignet, D. & Klemick, H. (2017), 'Modeling the Property Price Impact of Water Quality in 14 Chesapeake Bay Counties', *Ecological Economics* **135**, 103–113.

Walsh, P., Milon, J. W. & Scrogin, D. O. (2011), 'The Spatial Extent of Water Quality Benefits in Urban Housing Markets', *Land Economics* **87**(4), 628–644.

Wang, H. (2002), 'Pollution regulation and abatement efforts: Evidence from China', *Ecological Economics* **41**(1), 85–94.

Wang, H. & Wheeler, D. (2000), *Endogenous enforcement and effectiveness of China's pollution levy system*, Vol. 2336, World Bank Publications.

Wang, H. & Wheeler, D. (2003), 'Equilibrium pollution and economic development in China', *Environment and Development Economics* **8**(3), 451–466.

Wang, J. & Wei, Y. D. (2019), 'Agglomeration, environmental policies and surface water quality in china: A study based on a quasi-natural experiment', *Sustainability* **11**(19), 5394.

Wheeler, W. & Brown, M. J. (2013), 'Blood Lead Levels in Children Aged 1-5 Years – United States', *CDC Morbidity and Mortality Weekly Report* **62**(13), 245–248.

Willison, H. & Boyer, T. H. (2012), 'Secondary effects of anion exchange on chloride, sulfate, and lead release: Systems approach to corrosion control', *Water Research* **46**(7), 2385–2394.

Wilman, E. A. (1980), 'The value of time in recreation benefit studies', *Journal of Environmental Economics and Management* **7**(3), 272–286.

Wilson, J. K. & Damania, R. (2005), 'Corruption, political competition and environmental policy', *Journal of Environmental Economics and Management* **49**(3), 516–535.

Wolf, D. & Klaiber, H. A. (2017), 'Bloom and bust: Toxic algae's impact on nearby property values', *Ecological Economics* **135**, 209–221.

World Bank (2006), East Asia and Pacific environmental and social development discussion paper.

Xinhua Net (2018), 'National water quality monitoring network is established to ensure the accuracy of data (in chinese).'.
URL: *http://www.xinhuanet.com/local/2018-04/14/c_1122682188.htm*

Young, C. E. (1984), 'Perceived Water Quality And The Value Of Seasonal Homes', *JAWRA Journal of the American Water Resources Association* **20**, 163–166.

Zahran, S., Iverson, T., McElmurry, S. P. & Weiler, S. (2017), 'The effect of leaded aviation gasoline on blood lead in children', *Journal of the Association of Environmental and Resource Economists* **4**(2), 575–610.

Zaveri, E., Russ, J., Desbureaux, S., Damania, R., Rodella, A.-S. & Ribeiro, G. (2019), 'The nitrogen legacy'.

Zhang, B., Chen, X. & Guo, H. (2018), 'Does central supervision enhance local environmental enforcement ? Quasi-experimental evidence from China',

Journal of Public Economics **164**, 70–90.
URL: *https://doi.org/10.1016/j.jpubeco.2018.05.009*

Zhang, J. (2012), 'The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China', *Journal of Health Economics* **31**(1), 122–134.

Zhang, J. & Xu, L. C. (2016), 'The long-run effects of treated water on education: The rural drinking water program in China', *Journal of Development Economics* **122**, 1–15.

Zhang, Z., Gu, D. & Hayward, M. D. (2010), 'Childhood nutritional deprivation and cognitive impairment among older Chinese people', *Social Science & Medicine* pp. 1–9.

Zheng, H., Robinson, B. E., Liang, Y.-C., Polasky, S., Ma, D.-C., Wang, F.-C., Ruckelshaus, M., Ouyang, Z.-Y. & Daily, G. C. (2013), 'Benefits, costs, and livelihood implications of a regional payment for ecosystem service program', *Proceedings of the National Academy of Sciences* **110**(41), 16681–16686.

Zhou, R., Bi, X. & Segerson, K. (2020), 'Evaluating voluntary environmental programs with spillover effects', *Journal of the Association of Environmental and Resource Economists* **7**(1), 145–180.

Appendix D

Bibliography